

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2002, 18:46:43 : Search time 1824.97 Seconds  
(without alignments)  
6834.203 million cell updates/sec

Title: US-09-508-832-9  
Perfect score: 596  
Sequence: 1 atggcaagcaactctctga.....ctgggtggtggaatgcattg 596

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htgo\_inv.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID Query Description

RESULT	1	AX031287	AX031287	Sequence 9 from Patent WO9914321.	596 bp	DNA	Linear	PAT 20-SEP-2000
LOCUS		AX031287						
DEFINITION		AX031287						
ACCESSION		AX031287						
VERSION		AX031287.1	GI:10278618					
KEYWORDS								
SOURCE								
ORGANISM								
REFERENCE								
AUTHORS								
TITLE								
JOURNAL								
FEATURES								
Source								
CDS								

ALIGNMENTS

1	596	100.0	596	6	AX031287	Sequence
2	596	100.0	596	13	AX031313	Sequence
3	553	92.8	630	9	AF032457	Homo sapi
4	499.4	83.8	722	9	AB071197	Homo sapi
5	448.4	75.2	590	6	AX031283	Sequence
6	448.4	75.2	590	13	AX031309	Sequence
7	448.4	75.2	591	10	AF032459	Mus muscu
8	448.4	75.2	1137	10	BC019556	Mus muscu
9	443.6	74.4	591	10	AF065433	Rattus no
10	394.4	66.2	493	9	AB071198	Homo sapi
11	394.4	66.2	181497	9	AC096670	Homo sapi
12	294	49.3	416	6	AX031285	Sequence
13	294	49.3	416	13	AX031311	Sequence
14	207.6	34.8	422	6	AX031281	Sequence
15	207.6	34.8	422	13	AX031307	Sequence
16	207.6	34.8	423	10	AF032460	Mus muscu
17	207.6	34.8	423	10	AF136927	Rattus no
18	204.4	34.3	423	10	AF136927	Rattus no
19	201.4	33.8	443	9	AB071199	Homo sapi
20	197.4	33.1	542	9	AB071196	Homo sapi
21	197.4	33.1	542	10	AF065431	Rattus no
22	190.2	31.9	282	10	AB071200	Homo sapi
23	130.2	21.8	232	9	AB071200	Homo sapi
24	127	21.3	332	6	AX031279	Sequence
25	127	21.3	332	13	AX031305	Sequence
26	127	21.3	333	10	AF032461	Mus muscu
27	123.8	20.8	333	10	AF065432	Rattus no
28	108.8	18.3	2050	9	AF455755	Homo sapi
29	108.8	18.3	150035	30	AC046192	Homo sapi
30	108.8	18.3	166638	2	AC108463	Homo sapi
31	99	16.6	150035	30	AC046192	Homo sapi
32	65.4	11.0	59827	2	AC104222	Mus muscu
33	43	7.2	125020	9	AF429315	Homo sapi
34	39.2	6.6	181221	2	AC099532	Bos tauru
35	39.2	6.6	299350	1	SM5591786	Alsephen
36	37.8	6.3	770	9	HSPAX75	H. sapiens
37	37.4	6.3	120488	9	AC073257	Homo sapi
38	37.4	6.3	138119	2	AC021655	Homo sapi
39	37.4	6.3	151094	2	AC091797	Felis cat
40	37.4	6.3	171758	2	AC091761	Felis cat
41	37.4	6.3	172560	2	AC091794	Felis cat
42	37.4	6.3	185735	9	AC078785	Homo sapi
43	37.4	6.3	194006	9	AC074044	Homo sapi
44	37.4	6.3	204143	2	AC096461	Rattus no
45	37.4	6.3	231952	2	AC024108	Homo sapi

APP

O'Reilly, L., Puthalakath, H., Adams, J., O'Connor, L., Cory, S.,

Novel therapeutic molecules

Patent: WO 9914321-A 9 25-MAR-1999;  
INST MEDICAL W & E HALL (AU); PUTHALAKATH HANSA (AU); REILLY  
LORRANE O (AU); ADAMS JERRY (AU); CONNOR LIAM O (AU); CORY  
SUZANNE (AU); HUANG DAVID C S (AU); STRASSER ANDREAS (AU)

Location/Qualifiers

1..596  
/organism="unidentified"  
/db\_xref="taxon:32644"

/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC09657.1"  
/db\_xref="GI:10278619"  
/translation="MAKOPSDVSECDREGLOPAERPPOLRPGAPTSLOTPEQNP  
EGNHGGEDSCPHGSPGGLAPRSPGPRATSPPLIFMRRSSLLSRSSGYSFSDTD  
RSPAPMSDSTQTPSPPCQAFNHYLSAMASHROAPADMRPEIWIABELRIIGDEFN  
ATYARVFLNNYQAEEDHPRVILRLRYTVRLVWRMH"  
BASE COUNT 145 a 175 c 146 g 130 t  
ORIGIN

Query Match 100.0%; Score 596; DB 6; Length 596;  
Best Local Similarity 100.0%; Pred. No. 2.7e-149;  
Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 atggcaagcaaccccttcgatgaagttctgagttgacagagaagtagacaattgcag 60  
Db 1 ATGCCAAGCAACCTCTGATGAAGTCTGAGTGTGACGAGAGGTAGACAATTGCAG 60  
Oy 61 cctgcgagagggcctcccccagctcagacctggggccctacctcttctgctacagagcga 120  
Db 61 CCTGCGAGAGGCGCTCCCGACGCTCAGACCTGGGGCCCTACCTCCCTACAGACAGCCA 120  
Oy 121 caaggtcaatcctgaaggaatcagagaggtgaagggagacagctgccccacagcagccct 180  
Db 121 CAAGGTANTCTGAGGCAANTCAGGAGGTGAAGGGGACAGCTGCCCCACGCGACGCCCT 180  
Oy 181 caggcccgctggccccccctgcccagccctggcccttttctgctacagatccccgcttttc 240  
Db 181 CAGGGCCCGCTGGCCCCACCTGCCAGCCCTGGCCCTTTTCTACAGATCCCGGCTTTTC 240  
Oy 241 atctttatgagaagatcctccctcctgctcctcagatccctccagtggttatctctctttgac 300  
Db 241 ATCTTTATGAGAAGATCTCCCTGCTGCTCGATCTCCAGTGGGTATTTCTCTTTTAC 300  
Oy 301 acagacagagcccgagccacccatgagttgtgacaaatcaacacacacccagctccct 360  
Db 301 ACAGACAGAGCCCGAGCCACCATGATGTGTGACAAATCAACACAAACCCAGCTCCCT 360  
Oy 361 tgcagggcttcaacacatactcagtgcaatggcttccatgagggcgctgaacctgca 420  
Db 361 TGCAGGGCTTCAACACATATCTCAGTGCAATGGCTTCCATGAGGCGAGCTGAACCTGCA 420  
Oy 421 gatatgcgccagagatatgagatgcgccaaagagttgcggcgatcgatcgagagagtttaac 480  
Db 421 GATATGCCCGCAGAGATATGGATGCCCAAGAGTTGCGGCGTATCGGAGACGAGTTTAA 480  
Oy 481 gcttactatgcaagagaggtattttgaataattaccagagccgaagaccacacga 540  
Db 481 GCTTACTATGCAAGAGAGGATTTTGTGANTATTCAGACGCCCAAGCCACCCACCA 540  
Oy 541 atgggtatctacgactgttacgttacattgtccgctgggtgagagagatgcatg 596  
Db 541 ATGGTTATCTTACGACTGTTCAGTTACATTTGCGGCTGCTGTGGAGAAATGCATTC 596

RESULT 2  
AX031313  
LOCUS AX031313 596 bp DNA linear UNA 20-SEP-2000  
DEFINITION Sequence 9 from Patent WO9914321.  
ACCESSION AX031313  
VERSION AX031313.1 GI:10278641  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
1 (bases 1 to 596)  
O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,  
Huang,D.C. and Strasser,A.  
Novel therapeutic molecules  
Patent: WO 9914321-A 25-MAR-1999;  
JOURNAL INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HANSA (AU) ; REILLY

LOKRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY  
\* SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)  
FEATURES  
source  
Location/Qualifiers  
1..596  
/organism="unidentified"  
/db\_xref="taxon:32644"  
1..594  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC09662.1"  
/db\_xref="GI:10278642"  
/translation="MAKOPSDVSECDREGLOPAERPPOLRPGAPTSLOTPEQNP  
EGNHGGEDSCPHGSPGGLAPRSPGPRATSPPLIFMRRSSLLSRSSGYSFSDTD  
RSPAPMSDSTQTPSPPCQAFNHYLSAMASHROAPADMRPEIWIABELRIIGDEFN  
ATYARVFLNNYQAEEDHPRVILRLRYTVRLVWRMH"  
BASE COUNT 145 a 175 c 146 g 130 t  
ORIGIN

Query Match 100.0%; Score 596; DB 13; Length 596;  
Best Local Similarity 100.0%; Pred. No. 2.7e-149;  
Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 atggcaagcaaccccttcgatgaagttctgagttgacagagaagtagacaattgcag 60  
Db 1 ATGCCAAGCAACCTCTGATGAAGTCTGAGTGTGACGAGAGGTAGACAATTGCAG 60  
Oy 61 cctgcgagagggcctcccccagctcagacctggggccctacctcttctgctacagagcga 120  
Db 61 CCTGCGAGAGGCGCTCCCGACGCTCAGACCTGGGGCCCTACCTCCCTACAGACAGCCA 120  
Oy 121 caaggtcaatcctgaaggaatcagagaggtgaagggagacagctgccccacagcagccct 180  
Db 121 CAAGGTANTCTGAGGCAANTCAGGAGGTGAAGGGGACAGCTGCCCCACGCGACGCCCT 180  
Oy 181 caggcccgctggccccccctgcccagccctggcccttttctgctacagatccccgcttttc 240  
Db 181 CAGGGCCCGCTGGCCCCACCTGCCAGCCCTGGCCCTTTTCTACAGATCCCGGCTTTTC 240  
Oy 241 atctttatgagaagatcctccctcctgctcctcagatccctccagtggttatctctctttgac 300  
Db 241 ATCTTTATGAGAAGATCTCCCTGCTGCTCGATCTCCAGTGGGTATTTCTCTTTTAC 300  
Oy 301 acagacagagcccgagccacccatgagttgtgacaaatcaacacacacccagctccct 360  
Db 301 ACAGACAGAGCCCGAGCCACCATGATGTGTGACAAATCAACACAAACCCAGCTCCCT 360  
Oy 361 tgcagggccttcaacacatactcagtgcaatggcttccatgagggcgctgaacctgca 420  
Db 361 TGCAGGGCTTCAACACATATCTCAGTGCAATGGCTTCCATGAGGCGAGCTGAACCTGCA 420  
Oy 421 gatatgcgccagagatatgagatgcgccaaagagttgcggcgatcgatcgagagagtttaac 480  
Db 421 GATATGCCCGCAGAGATATGGATGCCCAAGAGTTGCGGCGTATCGGAGACGAGTTTAA 480  
Oy 481 gcttactatgcaagagaggtattttgaataattaccagagccgaagaccacacga 540  
Db 481 GCTTACTATGCAAGAGAGGATTTTGTGANTATTCAGACGCCCAAGCCACCCACCA 540  
Oy 541 atgggtatctacgactgttacgttacattgtccgctgggtgagagagatgcatg 596  
Db 541 ATGGTTATCTTACGACTGTTCAGTTACATTTGCGGCTGCTGTGGAGAAATGCATTC 596

RESULT 3  
AF032457  
LOCUS AF032457 597 bp mRNA linear PRI 19-FEB-1998  
DEFINITION Homo sapiens BimEL mRNA, complete cds.  
ACCESSION AF032457  
VERSION AF032457.1 GI:2895495  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens











```

QY 1 atggcaaaagcaaccttctgatgaagttcttgatgtgtgacagagaggtagacaaattgcag 60
DB 223 atggccaaagcaaccttctgatgaagttcttgatgtgtgacagagaggtagacaaattgcag 282
QY 61 cctgagagagggcctccacagtcagacctggggccctacccctccctacagagagacca 120
DB 283 cctgctcagagagggcctccacagtcagacctggggccctacccctccctacagagagacca 342
QY 121 caaggttaactcctgaagagcaatcacagagaggtgaaggggagagctgcccacccgagccct 180
DB 343 caaggttaactcctgaagagcaatcacagagaggtgaaggggagagctgcccacccgagccct 390
QY 181 cagggcccgctgacccacacacacacacacacacacacacacacacacacacacacacac 240
DB 391 cagggcccgctgacccacacacacacacacacacacacacacacacacacacacacacac 450
QY 241 atctttatgagagagatcctccctcctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300
DB 451 atctttatgagagagatcctccctcctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 510
QY 301 acagacagagggccgacacacacacacacacacacacacacacacacacacacacacacac 360
DB 511 acagacagagggccgacacacacacacacacacacacacacacacacacacacacacacac 570
QY 361 tgcagagcctcctcaacacacacacacacacacacacacacacacacacacacacacacac 414
DB 571 tgcagagcctcctcaacacacacacacacacacacacacacacacacacacacacacacac 630
QY 415 cctgagagatgcgcccagagacatgagtcgcccagagaggtgagggcgtatcggagagcag 474
DB 631 cctgagagatgcgcccagagacatgagtcgcccagagaggtgagggcgtatcggagagcag 690
QY 475 tttaacagagttatccttaagagagagagagagagagagagagagagagagagagagagac 534
DB 691 tttaacagagttatccttaagagagagagagagagagagagagagagagagagagagagac 750
QY 535 caacagagagttatccttaagagagagagagagagagagagagagagagagagagagagagac 594
DB 751 cctcaaatggttatcttacaactgtttacgcttttatcttccgctgtgtgtgtgtgtgtgt 810
QY 595 tg 596
DB 811 TG 812

RESULT 10
AF065433 591 bp mRNA linear ROD 11-MAR-1999
LOCUS Rattus norvegicus Bcl-2 related ovarian death gene product BOD-L
DEFINITION mRNA, complete cds.
ACCESSION AF065433
VERSION AF065433.1 GI:3228569
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 591)
Hsu, S.Y., Lin, P. and Hsueh, A.J.
BOD (Bcl-2-related ovarian death gene) is an ovarian BH3
domain-containing proapoptotic Bcl-2 protein capable of
dimerization with diverse antiapoptotic Bcl-2 members
Mol. Endocrinol. 12 (9), 1432-1440 (1998)
JOURNAL 98400436
MEDLINE 2 (bases 1 to 591)
AUTHORS Hsu, S.Y. and Hsueh, A.J.W.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-1998) GYM/OB, Stanford University, MSOB S385,
Stanford, CA 94305, USA
FEATURES
Location/Qualifiers
source 1..591

```

```

/organism="Rattus norvegicus"
/db_xref="taxon:10116"
1..591
/codon_start=1
/product="Bcl-2 related ovarian death gene product BOD-L"
/protein_id="AAC23595.1"
/db_xref="GI:3228570"
/translation="MAKOPSDVNSCEDREGGLOQPAERPPOLRPGAPTSIQTESQGNP
DGEDRCPHGSPQPLAPSPGPATPSPLFIFVRSSLLSKSSSYFSFOTDRSPA
PMSCDKSTQTPSPPCQAFNHYLASAMASIRQSEEPEDLRPEIRIAOELARRIGDEFNET
YTRRAFANDYREAEADHPQVILQLLRFLVWRRH"
BASE COUNT 139 a 177 c 153 g 122 t
ORIGIN

Query Match 74.48; Score 443.6; DB 10; Length 591;
Best Local Similarity 86.48; Pred. No. 2.1e-108;
Matches 520; Conservative 0; Mismatches 64; Indels 18; Gaps 2;

QY 1 atggcaaaagcaaccttctgatgaagttcttgatgtgtgacagagaggtagacaaattgcag 60
DB 1 atggccaaagcaaccttctgatgaagttcttgatgtgtgacagagaggtagacaaattgcag 60
QY 61 cctgagagagggcctccacagtcagacctggggccctacccctccctacagagagacca 120
DB 61 cctgctcagagagggcctccacagtcagacctggggccctacccctccctacagagagacca 120
QY 121 caaggttaactcctgaagagcaatcacagagaggtgaaggggagagctgcccacccgagccct 180
DB 121 caaggttaactcctgaagagcaatcacagagaggtgaaggggagagctgcccacccgagccct 180
QY 181 cagggcccgctgacccacacacacacacacacacacacacacacacacacacacacacac 240
DB 169 cagggcccgctgacccacacacacacacacacacacacacacacacacacacacacacac 228
QY 241 atctttatgagagagatcctccctcctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300
DB 229 atctttatgagagagatcctccctcctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 288
QY 301 acagacagagggccgacacacacacacacacacacacacacacacacacacacacacacac 360
DB 289 acagacagagggccgacacacacacacacacacacacacacacacacacacacacacacac 348
QY 361 tgcagagcctcctcaacacacacacacacacacacacacacacacacacacacacacacac 414
DB 349 tgcagagcctcctcaacacacacacacacacacacacacacacacacacacacacacacac 408
QY 415 cctgagagatgcgcccagagacatgagtcgcccagagaggtgagggcgtatcggagagcag 474
DB 409 cctgagagatgcgcccagagacatgagtcgcccagagaggtgagggcgtatcggagagcag 468
QY 475 tttaacagagttatccttaagagagagagagagagagagagagagagagagagagagagac 534
DB 469 tttaacagagttatccttaagagagagagagagagagagagagagagagagagagagagac 528
QY 535 caacagagagttatccttaagagagagagagagagagagagagagagagagagagagagagac 594
DB 529 cctcaaatggttatcttacaactgtttacgcttttatcttccgctgtgtgtgtgtgtgtgt 588
QY 595 tg 596
DB 589 TG 590

RESULT 11
AF071198 493 bp mRNA linear PRI 05-DEC-2001
LOCUS Homo sapiens mRNA for Bim-beta2, complete cds.
DEFINITION Homo sapiens mRNA for Bim-beta2, complete cds.
ACCESSION AF071198
VERSION AF071198.1 GI:17351902
KEYWORDS Homo sapiens cDNA to mRNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens

```

REFERENCE AC096670 AC013332  
 VERSION AC096670.1 GI:15668150  
 KEYWORDS HMG  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 U.M., Miyashita, T., Shikama, Y., Tadokoro, K. and Yamada, M.  
 Molecular cloning and characterization of six novel isoforms of  
 human Bim, a member of the proapoptotic Bcl-2 family(1)  
 FEBS Lett. 509 (1), 135-141 (2001)  
 11734221  
 2 (bases 1 to 493)  
 Miyashita, T.  
 Direct Submission  
 Submitted (06-SEP-2001) Toshiyuki Miyashita, National Children's  
 Medical Research Center, Department of Genetics; 3-35-31 Taishido,  
 Setagaya, Tokyo 154-8509, Japan (E-mail: tmiyashita@nch.go.jp,  
 Tel:81-3-3414-8121(ex.2763), Fax:81-3-3414-3208)  
 FEATURES  
 Location/Qualifiers  
 1..493  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 1..408  
 /gene="bim-beta2"  
 1..408  
 /gene="bim-beta2"  
 /note="alternatively spliced; see also Acc#  
 AB071195-AB071197, AB071199-AB071200"  
 /codon\_start=1  
 /product="Bim-beta2"  
 /protein\_id="F8A978592.1"  
 /db\_xref="GI:17351903."  
 /translation="MAKQSDVSSECDREGROLOPAERPPOLRPGATPSLOTPEQGNP  
 EGNHGGSCPHGSGOGPLAPSPGTPATSPILFIMRRSLLSRSSGYSFPTD  
 RSPAPMSCKSTQTPSPPCQAFNHYLSANGIFE"  
 119 a 151 c 114 g 109 t

## FEATURES

source

## gene

## CDS

Query Match 66.2%; Score 394.4; DB 9; Length 493;  
 Best Local Similarity 99.7%; Pred. No. 3.3e-95;  
 Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 atggcaaaacacctctgagtgaagttctgagtgagtgagcagaagtgagcaattgcag 60  
 Db 1 ATGGCAAAACACCTCTGAGTGAAGTCTGAGTGTGACCGAGAAGGTAGACAATTGCAG 60  
 Qy 61 cctgcggagagcctcccaagctcagacctgagcctgagcctcctcctacagacagagcca 120  
 Db 61 CCTGCGGAGAGGCTCCCGAGCTCAGACCTGGGGCCCTTACCCTCCCTACAGACAGAGCCA 120  
 Qy 121 caaggttaactcgaaggaatacagagagtgagtgagcagagctgccccacagagagccct 180  
 Db 121 CAAGGTAATCTGAAGGCAATCAGGAGGTGAGGGGACAGCTGCCCCCCACCGAGAGCCCT 180  
 Qy 181 caggcccgctggccccaccctgcagcctgagcctctttgtaccagatcccgcttttc 240  
 Db 181 CAGGGCCGCTGGCCCCACCTGCCAGCCCTGGCCCTTTTGTCTACCAAGATCCCGCTTTTC 240  
 Qy 241 atctttatgagagatccctcctgctgctcagatccctcagtggtgtattctctttgac 300  
 Db 241 ATCTTTATGAGAAGATCTCCTCCCTGCTCTCGATCCTCCAGTGGGTATTTCCTTTTTCAC 300  
 Qy 301 acagacagagagcccgagcccatgagttgtgacaaatcaacacaaaccccaagtcctcct 360  
 Db 301 ACAGACAGGAGCCCGAGCCCATGAGTTGTGACAAATCAACAAACCCCAAGTCTCTCCT 360  
 Qy 361 tgcaggccttcaaacactatctcagtgcaatggct 396  
 Db 361 TGCAGGCCCTTCAACCACTATCTCAGTGAATGGGT 396

## RESULT 12

AC096670/c

LOCUS

DEFINITION

Homo sapiens chromosome 2 clone RP11-438K19, complete sequence.  
 181497 bp DNA linear PRI 19-SEP-2001

## AC096670 AC013332

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

COMMENT

On Sep 19, 2001 this sequence version replaced gi:7248987.

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu/gsc/index.shtml

Contact: submissions@wustl.edu

----- Project Information -----

Center project name: H\_NH0438K19

Drafting center: WIBR

----- Location/Qualifiers -----

1..181497

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="2"

/clone="RP11-438K19"

BASE COUNT 52329 a 40542 c 39037 g 49529 t

ORIGIN

Query Match 66.2%; Score 394.4; DB 9; Length 181497;

Best Local Similarity 99.7%; Pred. No. 3.9e-95;

Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 atggcaaaacacctctgagtgaagttctgagtgagtgagcagaagtgagcaattgcag 60

Db 75475 ATGGCAAAACACCTCTGAGTGAAGTCTGAGTGTGACCGAGAAGGTAGACAATTGCAG 75416

Qy 61 cctgcggagagcctcccaagctcagacctgagcctgagcctcctcctacagacagagcca 120

Db 75415 CTTCCGAGAGAGCCCTCCCGAGCTCAGACCTGGGGCCCTTACCCTCCCTACAGACAGAGCCA 75356

Qy 121 caaggttaactcgaaggaatacagagagtgagtgagcagagctgccccacagagagccct 180

Db 75355 CAAGGTAATCTGAAGGCAATCAGGAGGTGAGGGGACAGCTGCCCCCCACCGAGAGCCCT 75296

Qy 181 caggcccgctggccccaccctgcagcctgagcctctttgtaccagatcccgcttttc 240

Db 75295 CAGGGCCGCTGGCCCCACCTGCCAGCCCTGGCCCTTTTGTCTACCAAGATCCCGCTTTTC 75236

Qy 241 atctttatgagagatccctcctgctgctcagatccctcagtggtgtattctctttgac 300

Db 75235 ATCTTTATGAGAAGATCTCCTCCCTGCTCTCGATCCTCCAGTGGGTATTTCCTTTTTCAC 75176

Qy 301 acagacagagagcccgagcccatgagttgtgacaaatcaacacaaaccccaagtcctcct 360

Db 75175 ACAGACAGGAGCCCGAGCCCATGAGTTGTGACAAATCAACAAACCCCAAGTCTCTCCT 75116

Qy 361 tgcaggccttcaaacactatctcagtgcaatggct 396

Db 75115 TGCAGGCCCTTCAACCACTATCTCAGTGAATGGGT 75080

RESULT 13

AC031285

LOCUS AX031285 416 bp DNA linear PAT 20-SEP-2000  
DEFINITION Sequence 7 from Patent WO9914321.  
ACCESSION AX031285  
VERSION AX031285.1 GI:10278616  
KEYWORDS  
SOURCE unclassified.  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 416)  
AUTHORS O'Reilly, L., Puthalakath, H., Adams, J., O'Connor, L., Cory, S., Huang, D.C. and Strasser, A.  
TITLE Novel therapeutic molecules  
JOURNAL Patent: WO 9914321-A 7 25-MAR-1999;  
INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HANSA (AU) ; REILLY LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)  
FEATURES  
source Location/Qualifiers  
1..416  
/organism="unclassified"  
/db\_xref="taxon:32644"  
1..>414  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC09656.1"  
/db\_xref="GI:10278617"  
/translation="MAKOPSDVSSDCREGROLQPAERPPQIRFGAPTSLQTEPDQRS PAPMSCKSTQTPSPPCOAFNHYLSAMSMRQAEADMRPEIMIAQELRRIGDEFNAY YARRVFLNNYQAEDHPRMVILRLRLVIVLRVHRH"  
BASE COUNT 113 a 113 c 103 g 87 t  
ORIGIN  
Query Match 49.3%; Score 294; DB 6; Length 416;  
Best Local Similarity 100.0%; Pred. No. 2.9e-68;  
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 303 agacagagccagaccatgagttgtgacaaatcaacaaaccccaagtcctcttg 362  
DB 123 AGACAGAGCCAGACCATGATGTCACAAATCAACAAACCCCAAGTCTCTTG 182  
QY 363 ccaggccttcacacatctatctcagtcgaatggtctccatgagcgagctgaacctgcaga 422  
DB 183 CCAGGCTTCAACCACTATCTCAGTGCATGCTTCCATGAGCGAGGCTGAACCTGCAGA 242  
QY 423 tatgcgcccagagataggtgacacaggttcgagcgatcgagcgagattcaacgc 482  
DB 243 TATGCGCCAGAGATATGTCGCCAAGAGTTCGCGCTATCGGAGACGAGTTTACGC 302  
QY 483 ttactatgcagaggggtatttttgaataattaccagcagcgagacccacacgaat 542  
DB 303 TTACTATGCAAGGAGGATATTTTGAATAATTACCAAGCCGAGACCCACCAAGAT 362  
QY 543 ggttattcagcacttaccgtatctcagtcgaatggtctccatgagcgagctgaacctgcaga 422  
DB 363 GGTATCTTACGACTGTACGTTACATGTCGCGCTGCTGCGAGATGCAATTG 416  
RESULT 15  
AF032458 417 bp mRNA linear PRI 19-FEB-1998  
LOCUS Homo sapiens BimL mRNA, complete cds.  
DEFINITION AF032458  
ACCESSION AF032458.1 GI:2895497  
VERSION  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 417)  
AUTHORS O'Connor, L., Strasser, A., O'Reilly, L.A., Hausmann, G., Adams, J.M., Cory, S. and Huang, D.C.  
TITLE Bim: a novel member of the Bcl-2 family that promotes apoptosis  
JOURNAL EMBO J. 17 (2), 384-395 (1998)  
MEDLINE 98094360  
PUBMED 9430630  
REFERENCE 2 (bases 1 to 417)  
AUTHORS Cory, S. and Huang, D.C.S.  
TITLE Direct Submission  
JOURNAL Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter & Eliza Hall Institute of Medical Research, PO Royal Melbourne Hospital, Parkville, Victoria 3050, Australia  
FEATURES  
source Location/Qualifiers  
1..417  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1..417  
CDS

LOCUS AX031285 416 bp DNA linear PAT 20-SEP-2000  
DEFINITION Sequence 7 from Patent WO9914321.  
ACCESSION AX031285  
VERSION AX031285.1 GI:10278616  
KEYWORDS  
SOURCE unclassified.  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 416)  
AUTHORS O'Reilly, L., Puthalakath, H., Adams, J., O'Connor, L., Cory, S., Huang, D.C. and Strasser, A.  
TITLE Novel therapeutic molecules  
JOURNAL Patent: WO 9914321-A 7 25-MAR-1999;  
INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HANSA (AU) ; REILLY LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)  
FEATURES  
source Location/Qualifiers  
1..416  
/organism="unclassified"  
/db\_xref="taxon:32644"  
1..>414  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC09661.1"  
/db\_xref="GI:10278640"  
/translation="MAKOPSDVSSDCREGROLQPAERPPQIRFGAPTSLQTEPDQRS PAPMSCKSTQTPSPPCOAFNHYLSAMSMRQAEADMRPEIMIAQELRRIGDEFNAY YARRVFLNNYQAEDHPRMVILRLRLVIVLRVHRH"  
BASE COUNT 113 a 113 c 103 g 87 t  
ORIGIN  
Query Match 49.3%; Score 294; DB 13; Length 416;  
Best Local Similarity 100.0%; Pred. No. 2.9e-68;  
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 303 agacagagccagaccatgagttgtgacaaatcaacaaaccccaagtcctcttg 362  
DB 123 AGACAGAGCCAGACCATGATGTCACAAATCAACAAACCCCAAGTCTCTTG 182  
QY 363 ccaggccttcacacatctatctcagtcgaatggtctccatgagcgagctgaacctgcaga 422  
DB 183 CCAGGCTTCAACCACTATCTCAGTGCATGCTTCCATGAGCGAGGCTGAACCTGCAGA 242  
QY 423 tatgcgcccagagataggtgacacaggttcgagcgatcgagcgagattcaacgc 482  
DB 243 TATGCGCCAGAGATATGTCGCCAAGAGTTCGCGCTATCGGAGACGAGTTTACGC 302  
QY 483 ttactatgcagaggggtatttttgaataattaccagcagcgagacccacacgaat 542  
DB 303 TTACTATGCAAGGAGGATATTTTGAATAATTACCAAGCCGAGACCCACCAAGAT 362  
QY 543 ggttattcagcacttaccgtatctcagtcgaatggtctccatgagcgagctgaacctgcaga 596  
DB 363 GGTATCTTACGACTGTACGTTACATGTCGCGCTGCTGCGAGATGCAATTG 416  
RESULT 15  
AF032458 417 bp mRNA linear PRI 19-FEB-1998  
LOCUS Homo sapiens BimL mRNA, complete cds.  
DEFINITION AF032458  
ACCESSION AF032458.1 GI:2895497  
VERSION  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 417)  
AUTHORS O'Connor, L., Strasser, A., O'Reilly, L.A., Hausmann, G., Adams, J.M., Cory, S. and Huang, D.C.  
TITLE Bim: a novel member of the Bcl-2 family that promotes apoptosis  
JOURNAL EMBO J. 17 (2), 384-395 (1998)  
MEDLINE 98094360  
PUBMED 9430630  
REFERENCE 2 (bases 1 to 417)  
AUTHORS Cory, S. and Huang, D.C.S.  
TITLE Direct Submission  
JOURNAL Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter & Eliza Hall Institute of Medical Research, PO Royal Melbourne Hospital, Parkville, Victoria 3050, Australia  
FEATURES  
source Location/Qualifiers  
1..417  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1..417  
CDS

```

/note="pro-apoptotic BH3-containing Bcl-2 family member".
/codon_start=1
/product="BimL"
/protein_id="AAC39594.1"
/db_xref="GI:2895498"
/translation="MKPQSVSSDCREGRLQPAERPPQLGAPTSLSQTEPDQRS
PAPMSCKSTQTPSPCAFNFHLSAMARQAPAPMPREITWIAELRRIGDEFNAY
YARVELNNYQAEDHPKRVILRLRLLYTVRLVWRMH"
a 113 c 103 g 87 t

```

BASE COUNT	114 a
ORIGIN	

Query Match	49.38;	Score 294;	DB 9;	Length 417;
Best Local Similarity	100.08;	Prod. NO. 2.9e-68;		
Matches 294;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	303	agacagagagccagcaccatgagttgtgcasaatcaacacaaaaccccaagctcctctg	362	
Db	123	AGACAGAGGCCAGCACCATGAGTTGTGACAANTCAACACAAACCCAAAGTCTCTTGG	182	
QY	363	caaggccttcaacacatctatctgaatggcttcctcctagggcaggctgaacctgcaga	422	
Db	183	CCAGGCTTCAACCACTATCTCAGTGCAAATGGCTTCATTGAGGAGGCTGAACCTGCAGA	242	
QY	423	tatgcgccagagatatgatatgcgccaaagagttgcggcgatacggagacagagtttaacgc	482	
Db	243	TATCGGCCACAGATATGATCGCCCAAGAGTTGCGGCTATCGGAGACAGGTTTAACGC	302	
QY	483	ttactatcgaaggaggtatttttgataataattcaacagcagccgaagaccaccacgaat	542	
Db	303	TTACTATGCAAGAGGGGTATTTTTGAATTAATTACCAAGCAGCGCAAGACCAACCCAGAA	362	
QY	543	ggttattctacgactgttacgttaccattgtccgcctgggtatggagaatgcattg	596	
Db	363	GGTATCTTACGACTCTGTACGTTACATGTTCGCCCTGGTGTGCAGAAATGCATTG	416	

Search completed: August 7, 2002, 20:11:14  
Job time: 5071 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2002, 19:36:18 ; Search time 210.72 Seconds  
(without alignments)  
4856.117 Million cell updates/sec

Title: US-09-508-832-9  
Perfect score: 596  
Sequence: 1 atgcacaaagcaaccttctga.....ctggtgacgaagatgcatg 596

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues  
Total number of hits satisfying chosen parameters: 3472872

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

```

Database :
N_Geneset_03802:**
1: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1980.DAT:**
2: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1981.DAT:**
3: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1982.DAT:**
4: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1983.DAT:**
5: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1984.DAT:**
6: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1985.DAT:**
7: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1986.DAT:**
8: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1987.DAT:**
9: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1988.DAT:**
10: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1989.DAT:**
11: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1990.DAT:**
12: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1991.DAT:**
13: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1992.DAT:**
14: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1993.DAT:**
15: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1994.DAT:**
16: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1995.DAT:**
17: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1996.DAT:**
18: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1997.DAT:**
19: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1998.DAT:**
20: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA1999.DAT:**
21: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA2000.DAT:**
22: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA2001A.DAT:**
23: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA2001B.DAT:**
24: /SIDSI/gcgdata/geneset/geneseqn-emb1/NA2002.DAT:**

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query a		Length	DB	ID	Description
		Match					
1	596	100.0		596	20	AAx24997	Human Bcl-2 intera
2	448.4	75.2		590	20	AAx24995	Murine Bcl-2 inter
3	294	49.3		416	20	AAx24996	Human Bcl-2 intera
4	276	46.3		442	23	AAx55350	DNA encoding novel
5	207.6	34.8		422	20	AAx24994	Murine Bcl-2 inter
6	127	21.3		332	20	AAx24993	Murine Bcl-2 inter
c 7	71	11.9		371	22	AA189284	Human polynucleoti
c 8	37	6.2		4047	21	AA338853	Human Jurkat cell
c 9	37	6.2		4047	21	AA238863	Human Jurkat cell

10	36.8	6.2	480	22	ABA42934	Human breast cell
11	36.8	6.2	480	22	ABA53355	Human foetal liver
12	36.8	6.2	480	22	ABA23135	Probe #1601 for ge
13	36.8	6.2	480	22	AAK01022	Human brain expres
14	36.8	6.2	480	22	AAK27064	Human bone marrow
15	36.8	6.2	480	22	AAI11862	Probe #1595 for ge
16	36.8	6.2	480	22	AAI12961	Probe #1647 used t
17	36.8	6.2	480	22	AAK01587	Human foetal liver
18	35.8	6.0	569	22	AAK63685	Human foetal liver
19	35.8	6.0	569	22	AAK30875	Probe #9341 for ge
20	35.8	6.0	569	22	AAK12201	Human brain expres
21	35.8	6.0	569	22	AAK37923	Human bone marrow
22	35.8	6.0	569	22	AAI18684	Probe #8617 for ge
23	35.8	6.0	569	22	AAI43797	Probe #12483 used
24	35.8	6.0	53522	22	AA513655	Genomic DNA sequen
25	35.6	6.0	16235	22	AAK66192	Human immune/haema
26	35.6	6.0	38186	20	AAK232028	Human METH1 relate
27	35.6	6.0	38186	22	AAK500085	Human METH1 relate
28	35.4	5.9	1128	21	AA259247	AC004449 cDNA clon
29	34.4	5.8	52616	22	AAK70459	Human NR8alpha/TPO
30	34.4	5.8	52616	22	AAK78930	Human immune/haema
31	33.4	5.6	5662	22	ABR19359	Human immune/haema
32	33.2	5.6	2567	22	ABR01379	Human nervous syst
33	33.2	5.6	3770	22	NAS01371	Human INTERCEPT 39
34	33	5.5	1464	16	AAQ76115	Human fetal cerebe
35	33	5.5	2913	16	AAQ76121	Human fetal cerebe
36	33	5.5	2923	16	AAQ76120	Human fetal cerebe
37	33	5.5	3183	16	AAQ76122	Human fetal brain
38	32.8	5.5	1942	22	AAI36808	Human musculoskele
39	32.6	5.5	786	23	AAK74228	DNA encoding novel
40	32.6	5.5	814	22	AAK49974	cDNA encoding nove
41	32.4	5.4	813	24	AA54817	Human DNA for pote
42	32.4	5.4	1806	23	AAK98127	Drosophila melanog
43	32.4	5.4	2238	23	AAK65829	DNA encoding novel
44	32.4	5.4	3328	22	AAI05135	Human reproductive
45	32.4	5.4	6196	23	ABL20074	Drosophila melanog

## ALIGNMENTS

RESULT	1
AAAX24997	
ID	AAAX24997 standard; cDNA; 596 bp.
XX	
XX	AAAX24997;
XX	
XX	05-JUL-1999 (first entry)
XX	
DE	Human Bcl-2 interacting mediator of cell death Bim-EL cDNA.
XX	
XX	Bim-EL; Bcl-2 interacting mediator of cell death; apoptosis;
KW	cell cycle; human; cancer; autoimmune disease;
KW	degenerative disease; therapy; contraceptive; splice variant;
KW	isoform; ss.

DR P-PSDB: AAW98158.  
 XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer  
 PT treatment  
 XX  
 PS Claim 7: Page 101-102; 145pp; English.  
 XX The present sequence encodes the extra long form (EL) of human Bim,  
 CC or Bcl-2 interacting mediator of cell death (see AAW98158), a novel  
 CC member of the Bcl-2 family that is capable of inducing cell death  
 CC (apoptosis) and which acts as a 'death-ligand' for certain members  
 CC of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the  
 CC only Bcl-2 homology region which it encompasses is BH3. It is the  
 CC only BH3-only protein for which splice variants exist. These  
 CC result in the expression of a variety of isoforms, i.e. Bim-S,  
 CC Bim-L and Bim-EL. cDNAs encoding human Bim-L and Bim-EL (see  
 CC AAW98158) were isolated from embryo and liver cDNA libraries using  
 CC mouse b1m cDNA. Murine Bim-S, Bim-L and Bim-EL isoforms (see  
 CC AAW98154-56) are also provided. The human Bim gene maps to  
 CC chromosome 2 at bands 2q12-2q13. Binding the dynein light  
 CC chain was shown to regulate the pro-apoptotic activity of Bim.  
 CC Bim-S, the splice variant which does not bind to dynein light  
 CC chain, is a much more potent killer than either Bim-L or Bim-EL.  
 CC The invention provides variants (see AAW98159-68) of murine and human  
 CC Bim-L or Bim-EL that cannot bind, couple or otherwise associate  
 CC with a dynein light chain. The identification of Bim permits the  
 CC identification and rational design of a range of products for use  
 CC in therapy, diagnosis, antibody generation and involving modulation  
 CC of physiological cell death. These therapeutic molecules may act  
 CC as either antagonists or agonists of Bim's function and will be  
 CC useful in cancer, autoimmune or degenerative disease therapy.  
 CC Increased Bim expression or Bim activity is useful, e.g. for  
 CC treatment or prophylaxis in conditions such as cancer and deletion  
 CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim  
 CC expression of Bim activity is useful in regulating inhibition or  
 CC prevention of cell death or degeneration and chemotherapy or during  
 CC conditions during e.g. gamma-irradiation and chemotherapy or during  
 CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,  
 CC hypoxia, degenerative diseases or for prolonging the survival of  
 CC cells being transplanted for treatment of disease. Since Bim is  
 CC expressed in germ cells, modulating Bim expression or Bim activity  
 CC is useful, e.g. as a contraceptive or method of sterilization by  
 CC preventing generation of fertile sperm.  
 XX  
 SQ Sequence 596 BP; 145 A; 175 C; 146 G; 130 T; 0 other;

Query Match 100.0%; Score 596; DB 20; Length 596;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-178;  
 Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atggcaagcaacctctgatgtaagtctgagtgacgagaggtagacaattgcag 60  
 Db 1 atggcaagcaacctctgatgtaagtctgagtgacgagaggtagacaattgcag 60  
 Qy 61 ctgcgagagagcctcccaagctcagacctgggcccctacctccctacagacagacca 120  
 Db 61 ctgcgagagagcctcccaagctcagacctgggcccctacctccctacagacagacca 120  
 Qy 121 caaggttaacctgaaggcaatcacggagggtgaaggagagctgcctcccaagagccct 180  
 Db 121 caaggttaacctgaaggcaatcacggagggtgaaggagagctgcctcccaagagccct 180  
 Qy 181 caggggccctgagcccccacccgagccctggccctttgtctaccagatcccccgctttc 240  
 Db 181 caggggccctgagcccccacccgagccctggccctttgtctaccagatcccccgctttc 240  
 Qy 241 attcttatgagaagatcctcctctgctctcagatcctccagtggtattctctttgac 300  
 Db 241 attcttatgagaagatcctcctctgctctcagatcctccagtggtattctctttgac 300  
 Qy 301 acagacagagagcccaagcccatgattgtgacaaatcaacacaccccaagctcctct 360  
 Db 301 acagacagagagcccaagcccatgattgtgacaaatcaacacaccccaagctcctct 360

Db 301 acagacagagagcccaagcccatgattgtgacaaatcaacacaccccaagctcctct 360  
 Qy 361 tgcagagccttcaaacactatctcagtgcaatgctctccatgagcagctgaactgca 420  
 Db 361 tgcagagccttcaaacactatctcagtgcaatgctctccatgagcagctgaactgca 420  
 Qy 421 gatatgcgcccagagatgatgagtcgcccagagttgcggcgatcgcgagacgagtttaac 480  
 Db 421 gatatgcgcccagagatgatgagtcgcccagagttgcggcgatcgcgagacgagtttaac 480  
 Qy 481 gcttactatgcagagaggggtatttttgataattaccacagcagcccaagacccaccacga 540  
 Db 481 gcttactatgcagagaggggtatttttgataattaccacagcagcccaagacccaccacga 540  
 Qy 541 atggttatcttaccagctgttaagttacatttcgctgggtgagagaatgcattg 596  
 Db 541 atggttatcttaccagctgttaagttacatttcgctgggtgagagaatgcattg 596

RESULT 2  
 AAX24995  
 ID AAX24995 standard; cDNA; 590 BP.  
 XX AC AAX24995;  
 XX 05-JUL-1999 (first entry)  
 DT Murine Bcl-2 interacting mediator of cell death Bim-EL cDNA.  
 DE Bim-EL; Bcl-2 interacting mediator of cell death; apoptosis;  
 KW cell cycle; mouse; cancer; autoimmune disease;  
 KW degenerative disease; therapy; contraceptive; splice variant;  
 KW isoform; ss.  
 XX OS Mus musculus.  
 XX WO9914321-A1.  
 XX 25-MAR-1999.  
 XX 17-SEP-1998; 98WO-AU00772.  
 XX 24-SEP-1997; 97AU-0009373.  
 XX 17-SEP-1997; 97AU-0009263.  
 XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
 XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;  
 PI Puthalakath H, Strasser A;  
 XX WP1; 1999-244030/20.  
 DR P-PSDB; AAW98156.  
 PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer  
 PT treatment  
 XX Claim 3; Page 96-97; 145pp; English.

XX The present sequence encodes the extra long form (EL) of murine Bim,  
 CC or Bcl-2 interacting mediator of cell death (see AAW98156), a novel  
 CC member of the Bcl-2 family that is capable of inducing cell death  
 CC (apoptosis) and which acts as a 'death-ligand' for certain members  
 CC of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the  
 CC only Bcl-2 homology region which it encompasses is BH3. It is the  
 CC only BH3-only protein for which splice variants exist. These  
 CC result in the expression of a variety of isoforms, i.e. Bim-S,  
 CC Bim-L and Bim-EL (see AAW98154-56). cDNAs encoding these murine Bim  
 CC isoforms were obtained from a T lymphoma cDNA library using human  
 CC recombinant Bcl-2 protein. The murine Bim gene has been mapped to  
 CC chromosome 2 at bands F3-G. Human Bim-L and Bim-EL isoforms have  
 CC also been identified (see AAW98157-58). Binding the dynein light  
 CC chain was shown to regulate the pro-apoptotic activity of Bim.  
 CC Bim-S, the splice variant which does not bind to dynein light



CC chain, is a much more potent killer than either Bim-L or Bim-EL. CC  
CC The invention provides variants (see AA98159-68) of murine and human CC  
CC Bim-L or Bim-EL that cannot bind, couple or otherwise associate CC  
CC with a dynein light chain. The identification of Bim permits the CC  
CC identification and rational design of a range of products for use CC  
CC in therapy, diagnosis, antibody generation and involving modulation CC  
CC of physiological cell death. These therapeutic molecules may act CC  
CC as either antagonists or agonists of Bim's function and will be CC  
CC useful in cancer, autoimmune or degenerative disease therapy. CC  
CC Increased Bim expression or Bim activity is useful, e.g. for CC  
CC treatment or prophylaxis in conditions such as cancer and deletion CC  
CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim CC  
CC expression of Bim activity is useful in regulating inhibition or CC  
CC prevention of cell death or degeneration such as under cytotoxic CC  
CC conditions during e.g. gamma-irradiation and chemotherapy or during CC  
CC HIV/AIDS or other viral infections, ischemia, myocardial infarction, CC  
CC hypoxia, degenerative diseases or for prolonging the survival of CC  
CC cells being transplanted for treatment of disease. Since Bim is CC  
CC expressed in germ cells, modulating Bim expression or Bim activity CC  
CC is useful, e.g. as a contraceptive or method of sterilization by CC  
CC preventing generation of fertile sperm.

xx Sequence 590 BP; 137 A; 178 C; 150 G; 125 T; 0 other;  
SQ

Query Match 75.2%; Score 448.4; DB 20; Length 590;  
Best Local Similarity 86.9%; Pred. No. 1.9e-131;  
Matches 523; Conservative 0; Mismatches 61; Indels 18; Gaps 2;

Qy	1	atggcaagcaactctctgatgtaaagtctgtgacccgagaggttagacaattgcag	60
Db			
Db	1	atggccaagcaactctctgatgtaaagtctgtgacagaggttgacaattgcag	60
Qy	61	ctctggagagcctcccaagctcagacctgggcccctacctccacagacagacca	120
Db			
Db	61	ctctgtgagagcctcccaagctcaggcctgggcccctacctccacagacagaacg	120
Qy	121	caagctaatctgaagcaatcacggaggtgaaggaggaagctccccacagcagccct	180
Db			
Db	121	caaggtaatctccagaggc-----gaaggggaccgctgcccaccagggagccct	168
Qy	181	cagggcccgcctggcccacactgcagccctggcccttttgtaccagatccccgccttc	240
Db			
Db	169	caaggcccgcctggcccacccggcagccctggccctttgtaccagatccccacttctc	228
Qy	241	attctttatgagaagatctctcctgctgtctcagctccagctgggtattctcttttgac	300
Db			
Db	229	attctttgtgagaagatctctcctgctgtcccggtccctcagtggggtattctcttttgac	288
Qy	301	acagacaggagcccgaccacccaatgagttgtgacaaatcaacacaaacccaagtctcct	360
Db			
Db	289	acagacaggagcccgaccacccaatgagttgtgacaagtcaacacaaacccaagtctcct	348
Qy	361	tgcaggccttcaaccactatctcagtgcaatggcttccatgaggcagcgt-----gaa	414
Db			
Db	349	tgcaggccttcaaccactatctcagtgcaatggcttccatagcagcagctcagaggagaa	408
Qy	415	cttgacagatatgcgccagagatatggatgcccgaagagtgcggcgctatcgagacagag	474
Db			
Db	409	cttgaaagatctgcccggagatcacggatgtgcacagagagctgcgcggagctcgagacag	468
Qy	475	tttaacgccttaactgcaagagggtatttttgaataattaccagagcagccggaagaccac	534
Db			
Db	469	ttcaacgaaactacacaaggagggtgtttgtcaaatgattaccgcgaggtgtaagaccac	528
Qy	535	ccacgaatgggtattcttaacagctgttacgtttacattgttcgcgcctgggtgaggaatgcac	594
Db			
Db	529	ctctcaaatgggtattcttaacagctgttacgtttacattgttcgcgcctgggtgaggaatgcac	588
Qy	595	tg	596
Db			
Db	589	tg	590

CC expressed in germ cells, modulating Bim expression or Bim activity  
 CC is useful, e.g. as a contraceptive or method of sterilization by  
 CC preventing generation of fertile sperm.  
 XX  
 SQ Sequence 416 BP; 113 A; 113 C; 103 G; 87 T; 0 other;

Query Match 49.3%; Score 294; DB 20; Length 416;  
 Best Local Similarity 100.0%; Pred. No. 1e-82;  
 Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 303 agacagagccagcaccatggtgtgacaaatcaacacacaccccaagctctcttg 362  
 DB 123 agacagagccagcaccatggtgtgacaaatcaacacacaccccaagctctcttg 182  
 OY 363 ccaggcttcaaccatctcagtgcaatgggttccatgagcgaggtgaacctgcaga 422  
 DB 183 ccaggcttcaaccatctcagtgcaatgggttccatgagcgaggtgaacctgcaga 242  
 OY 423 tatgcgccccagagatggtgacgccccagaggttgccggtatcgagacgagtttaacgc 482  
 DB 243 tatgcgccccagagatggtgacgccccagaggttgccggtatcgagacgagtttaacgc 302  
 OY 483 ttactatgcaagagggatattttgaaataattaccagcagcgaagaccacacgaat 542  
 DB 303 ttactatgcaagagggatattttgaaataattaccagcagcgaagaccacacgaat 362  
 OY 543 ggttatcttaccagctgttaccgttaccattgtccgctgggtggaagatgcattg 596  
 DB 363 ggttatcttaccagctgttaccgttaccattgtccgctgggtggaagatgcattg 416

## RESULT 4

AAS65350  
 ID AAS65350 standard; cDNA: 442 BP.

XX  
 AC AAS65350;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #1154.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; sa.  
 XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR P-PSDB; ABG01163.

XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -

XX Claim 1; SEQ ID No 1154; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 442 BP; 94 A; 154 C; 105 G; 89 T; 0 other;

Query Match 46.3%; Score 276; DB 23; Length 442;  
 Best Local Similarity 100.0%; Pred. No. 5.1e-77;

Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atgcaaaagcaacctctctgatgaagttctgagtgtagcagagaagtagacaattgcag 60  
 DB 167 atgcaaaagcaacctctctgatgaagttctgagtgtagcagagaagtagacaattgcag 226  
 OY 61 cctgcggagagggctcccccagctcagacctggggccctacctccctcacagacagacca 120  
 DB 227 cctgcggagagggctcccccagctcagacctggggccctacctccctcacagacagacca 286  
 OY 121 caaggttaactgaaggcaatcacgaggtgaaggagcagctgccccacggcagacct 180  
 DB 287 caaggttaactgaaggcaatcacgaggtgaaggagcagctgccccacggcagacct 346  
 OY 181 caggggcccgctggccccacctgcccagccctggcccttttctaccagatcccccgcttttc 240  
 DB 347 caggggcccgctggccccacctgcccagccctggcccttttctaccagatcccccgcttttc 406  
 OY 241 acctttatgagaagatccctccctgtgtctcgatcc 276  
 DB 407 acctttatgagaagatccctccctgtgtctcgatcc 442

## RESULT 5

AAX24994  
 ID AAX24994 standard; cDNA: 422 BP.

XX  
 AC AAX24994;

XX 05-JUL-1999 (first entry)

DE Murine Bcl-2 interacting mediator of cell death Bim-L cDNA.

XX Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;  
 KW cell cycle; mouse; cancer; autoimmune disease;  
 KW degenerative disease; therapy; contraceptive; splice variant;  
 KW isoform; ss.

XX Mus musculus.

XX WO9914321-A1.

XX 25-MAR-1999.

XX 17-SEP-1998; 98WO-AU00772.

XX 24-SEP-1997; 97AU-0009373.

XX 17-SEP-1997; 97AU-0009263.

PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
 XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;  
 PI Puthalakath H, Strasser A;  
 PI WPI: 1999-244030/20.  
 DR P-PSDB: AAW98155.  
 XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer  
 PT treatment  
 PT Claim 3; Page 94-95; 145pp; English.  
 PS The present sequence encodes the long form (L) of murine Bim, or  
 XX Bcl-2 interacting mediator of cell death (see AAW98155), a novel  
 CC member of the Bcl-2 family that is capable of inducing cell death  
 CC (apoptosis) and which acts as a 'death-ligand' for certain members  
 CC of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the  
 CC only Bcl-2 homology region which it encompasses is BH3. It is the  
 CC only BH3-only protein for which splice variants exist. These  
 CC result in the expression of a variety of isoforms, i.e. Bim-S,  
 CC Bim-L and Bim-EL (see AAW98154-56). cDNAs encoding these murine Bim  
 CC isoforms were obtained from a T lymphoma cDNA library using human  
 CC recombinant Bcl-2 protein. The murine Bim gene has been mapped to  
 CC chromosome 2 at bands F3-G. Human Bim-L and Bim-EL isoforms have  
 CC also been identified (see AAW98157-58). Binding the dynein light  
 CC chain was shown to regulate the pro-apoptotic activity of Bim.  
 CC Bim-S, the splice variant which does not bind to dynein light  
 CC chain, is a much more potent killer than either Bim-L or Bim-EL.  
 CC The invention provides variants (see AAW98159-68) of murine and human  
 CC Bim-L or Bim-EL that cannot bind, couple or otherwise associate  
 CC with a dynein light chain. The identification of Bim permits the  
 CC identification and rational design of a range of products for use  
 CC in therapy, diagnosis, antibody generation and involving modulation  
 CC of physiological cell death. These therapeutic molecules may act  
 CC as either antagonists or agonists of Bim's function and will be  
 CC useful in cancer, autoimmune or degenerative disease therapy.  
 CC Increased Bim expression or Bim activity is useful, e.g. for  
 CC treatment or prophylaxis in conditions such as cancer and deletion  
 CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim  
 CC expression of Bim activity is useful in regulating inhibition or  
 CC prevention of cell death or degeneration such as under cytotoxic  
 CC conditions during e.g. gamma-irradiation and chemotherapy or during  
 CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,  
 CC hypoxia, degenerative diseases or for prolonging the survival of  
 CC cells being transplanted for treatment of disease. Since Bim is  
 CC expressed in germ cells, modulating Bim expression or Bim activity  
 CC is useful, e.g. as a contraceptive or method of sterilization by  
 CC preventing generation of fertile sperm.  
 XX Sequence 422 BP; 112 A; 116 C; 109 G; 85 T; 0 other;  
 50

Query Match 34.8%; Score 207.6; DB 20; Length 422;  
 Best Local Similarity 83.3%; Pred. No. 2.1e-55;  
 Matches 250; Conservative 0; Mismatches 44; Indels 6; Gaps 1;

Qy 303 agacagagccagccatcagtgatggtgacaaatcaacacaccccaagtcctctg 362  
 Db 123 agcagagccgcccagccatcagtgatggtgacaaatcaacacaccccaagtcctctg 182

Qy 363 ccaggccttcaacactatcagtgcaatggtcctccatgagcgagct-----gaacc 416  
 Db 183 ccaggccttcaacactatcagtgcaatggtcctccatgagcgagcgagcacc 242

Qy 417 tcagatatacgccagagatagatgagtcgcccagagtggtcgcgatcgagagcaggtt 476  
 Db 243 tgaagactgcccgcagagatgagtcgcccagagtggtcgcgatcgagagcaggtt 302

Qy 477 taacgcttactgcaagaggggtatttttgaataattaccacagcagcagacaccc 536  
 Db 303 caacgaacttacagaggggtgtttgcaaatgattaccgcgaggtgagacaccc 362

Qy 537 acgaatggtatctctacgaactgttacgttacattgtccgctggtggtggaagatgcattg 596  
 Db 363 tcaaatggtatctctacgaactgttacgttacattgtccgctggtggtggaagatgcattg 422

RESULT 6  
 AAX24993  
 ID AAX24993 standard; CDNA; 332 BP.  
 XX  
 AC AAX24993;  
 XX  
 DT 05-JUL-1999 (first entry)  
 XX  
 DE Murine Bcl-2 interacting mediator of cell death Bim-S cDNA.  
 XX  
 KW Bim-S; Bcl-2 interacting mediator of cell death; apoptosis;  
 KW cell cycle; mouse; cancer; autoimmune disease;  
 KW degenerative disease; therapy; contraceptive; splice variant;  
 KW isoform; ss.  
 XX  
 OS Mus musculus.  
 XX  
 PN W09914321-A1.  
 XX  
 PD 25-MAR-1999.  
 XX  
 PF 17-SEP-1998; 98WO-AU00772.  
 XX  
 PR 24-SEP-1997; 97AU-0009373.  
 PR 17-SEP-1997; 97AU-0009263.  
 XX  
 PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
 PI Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;  
 PI Puthalakath H, Strasser A;  
 XX WPI: 1999-244030/20.  
 DR P-PSDB: AAW98154.  
 XX  
 PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer  
 PT treatment  
 PT Claim 3; Page 92; 145pp; English.  
 PS The present sequence encodes the short form (S) of murine Bim, or  
 CC Bcl-2 interacting mediator of cell death (see AAW98154), a novel  
 CC member of the Bcl-2 family that is capable of inducing cell death  
 CC (apoptosis) and which acts as a 'death-ligand' for certain members  
 CC of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the  
 CC only Bcl-2 homology region which it encompasses is BH3. It is the  
 CC only BH3-only protein for which splice variants exist. These  
 CC result in the expression of a variety of isoforms, i.e. Bim-S,  
 CC Bim-L and Bim-EL (see AAW98154-56). cDNAs encoding these murine Bim  
 CC isoforms were obtained from a T lymphoma cDNA library using human  
 CC recombinant Bcl-2 protein. The murine Bim gene has been mapped to  
 CC chromosome 2 at bands F3-G. Human Bim-L and Bim-EL isoforms have  
 CC also been identified (see AAW98157-58). Binding the dynein light  
 CC chain was shown to regulate the pro-apoptotic activity of Bim.  
 CC Bim-S, the splice variant which does not bind to dynein light  
 CC chain, is a much more potent killer than either Bim-L or Bim-EL.  
 CC The invention provides variants (see AAW98159-68) of murine and human  
 CC Bim-L or Bim-EL that cannot bind, couple or otherwise associate  
 CC with a dynein light chain. The identification of Bim permits the  
 CC identification and rational design of a range of products for use  
 CC in therapy, diagnosis, antibody generation and involving modulation  
 CC of physiological cell death. These therapeutic molecules may act  
 CC as either antagonists or agonists of Bim's function and will be  
 CC useful in cancer, autoimmune or degenerative disease therapy.  
 CC Increased Bim expression or Bim activity is useful, e.g. for  
 CC treatment or prophylaxis in conditions such as cancer and deletion  
 CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim  
 CC expression of Bim activity is useful in regulating inhibition or  
 CC prevention of cell death or degeneration such as under cytotoxic





CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human breast and Bt 474 cells. The method involves contacting  
 CC the probes with a collection of detectably labelled nucleic acids  
 CC derived from mRNA of human breast, and then measuring the label  
 CC bound to each probe of the microarray. The probes are useful for  
 CC verifying the expression of regions of genomic DNA predicted to  
 CC encode proteins. They are useful for gene discovery, and for  
 CC determining predisposition and/or prognosing breast disease. Gene  
 CC expression analysis is useful for assessing the toxicity of chemical  
 CC agents on cells. The microarray of this invention presents a far greater  
 CC diversity of probes for measuring gene expression, with far less bias  
 CC than expressed sequence tag microarrays. The method is suitable for  
 CC rapid production of functional information from genomic sequences. The  
 CC present sequence is a single exon nucleic acid probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 480 BP; 109 A; 162 C; 129 G; 80 T; 0 other;

Query Match 6.2%; Score 36.8; DB 22; Length 480;  
 Best Local Similarity 50.0%; Pred. No. 0.21; Indels 0; Gaps 0;  
 Matches 92; Conservative 0; Mismatches 0; Gaps 0;  
 QY 65 cggagagcctcccccagctcagacctggggccctaccctccctacagacagaccacaag 124  
 DB 22 cgcctggacttgaccaccagcagcagtgcatgacgcccagctggatgcaggtgcacaga 81  
 QY 125 gtaactctgaaggcaatcacggagtggaaggacagctgccccacggcagcctcagg 184  
 DB 82 tgatcccgccagctcttcactcaggaggagcagcagtgggcccccagtcacacagctcccg 141  
 QY 185 gccctggcccccagctcagacctggcccttttcttctacagatccccgcttttcatct 244  
 DB 142 gctctcccccacccacacatcgtggggaaatttgctctccaggccacagcttctcttc 201  
 QY 245 ttat 248  
 DB 202 atat 205

RESULT 11  
 ABA53355  
 ID ABA53355 standard; DNA; 480 BP.  
 AC ABA53355;  
 XX  
 DT 01-FEB-2002 (first entry)  
 XX  
 DE Human foetal liver single exon nucleic acid probe #1660.  
 XX  
 KW Human: foetal liver; gene expression; single exon nucleic acid probe; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157277-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00669.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-483447/52.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human fetal liver -  
 XX  
 PS Claim 1; SEQ ID NO 1660; 639pp + sequence listing; English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC fetal liver. The present sequence is a single exon nucleic acid  
 CC probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 480 BP; 109 A; 162 C; 129 G; 80 T; 0 other;

Query Match 6.2%; Score 36.8; DB 22; Length 480;  
 Best Local Similarity 50.0%; Pred. No. 0.21; Indels 0; Gaps 0;  
 Matches 92; Conservative 0; Mismatches 0; Gaps 0;  
 QY 65 cggagagcctcccccagctcagacctggggccctaccctccctacagacagaccacaag 124  
 DB 22 cgcctggacttgaccaccagcagcagtgcatgacgcccagctggatgcaggtgcacaga 81  
 QY 125 gtaactctgaaggcaatcacggagtggaaggacagctgccccacggcagcctcagg 184  
 DB 82 tgatcccgccagctcttcactcaggaggagcagcagtgggcccccagtcacacagctcccg 141  
 QY 185 gccctggcccccagctcagacctggcccttttcttctacagatccccgcttttcatct 244  
 DB 142 gctctcccccacccacacatcgtggggaaatttgctctccaggccacagcttctcttc 201  
 QY 245 ttat 248  
 DB 202 atat 205

RESULT 12  
 ABA23135  
 ID ABA23135 standard; DNA; 480 BP.  
 XX  
 AC ABA23135;  
 XX  
 DT 23-JAN-2002 (first entry)  
 XX  
 DE Probe #1601 for gene expression analysis in human heart cell sample.  
 XX  
 KW Human; gene expression; heart; microarray; vascular system; probe;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157274-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00666.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX



```

PR 21-SEP-2000: 2000US-0234687.
PR 27-SEP-2000: 2000US-0236359.
PR 04-OCT-2000: 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI: 2001-488900/53.
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow.
XX Example 4; SEQ ID NO: 1621; 658pp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukemia and myeloma. The present sequence is one of
XX the probes of the invention.
XX
XX Sequence 480 BP; 109 A; 162 C; 129 G; 80 T; 0 other;
SQ

```

Query Match 5.2%; Score 36.8; DB 22; Length 480;  
 Best Local Similarity 50.0%; Pred. No. 0.21;  
 Matches 92; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

```

Qy 65 cggagagggctcccgctcagacgtggggccctacccctccctacagagagccacaag 124
Db 22 cgcgtggacttgaccaccagagcagtgcatgcccagctggtgagtgacaga 81
Qy 125 gtaactctgaagcaatcacggaggtgaaggagcagctgcccacggcagcctcagg 184
Db 82 tgatccgcagctcttcacctagggagcagcagctggcccagtgccagcgtcccg 141
Qy 185 gccgcgtggcccaactcagacccctggcccttttctacacagatccccgctttctat 244
Db 142 gccgtctccaccaccacacatcgtggggaattggtctccaggtcacaggttctcttc 201
Qy 245 ttat 248
Db 202 atat 205

```

RESULT 15  
 AA11662  
 ID AA11662 standard; DNA; 480 BP.  
 AC AA11662;  
 DT 12-OCT-2001 (first entry)  
 DE Probe #1595 for gene expression analysis in human cervical cell sample.  
 KW Probe: human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer; 86.  
 OS Homo sapiens.  
 PN WO200157278-A2.  
 PD 09-AUG-2001.  
 XX 30-JAN-2001; 2001WO-US00670.  
 XX 04-FEB-2000; 2000US-0180312.  
 XX 26-MAY-2000; 2000US-0207456.  
 XX 30-JUN-2000; 2000US-0608408.  
 XX 03-AUG-2000; 2000US-0632366.  
 XX 21-SEP-2000; 2000US-0234687.  
 XX 27-SEP-2000; 2000US-0236359.

```

PR 04-OCT-2000: 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI: 2001-488901/53.
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells.
XX Claim 25; SEQ ID NO 1595; 487pp; English.
XX The present invention relates to human single exon nucleic acid probes
XX (SENPs). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging
XX of diseases of the cervix, notably cervical cancer.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 480 BP; 109 A; 162 C; 129 G; 80 T; 0 other;
SQ

```

Query Match 6.2%; Score 36.8; DB 22; Length 480;  
 Best Local Similarity 50.0%; Pred. No. 0.21;  
 Matches 92; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

```

Qy 65 cggagagggctcccgctcagacgtggggccctacccctccctacagagagccacaag 124
Db 22 cgcgtggacttgaccaccagagcagtgcatgcccagctggtgagtgacaga 81
Qy 125 gtaactctgaagcaatcacggaggtgaaggagcagctgcccacggcagcctcagg 184
Db 82 tgatccgcagctcttcacctagggagcagcagctggcccagtgccagcgtcccg 141
Qy 185 gccgcgtggcccaactcagacccctggcccttttctacacagatccccgctttctat 244
Db 142 gccgtctccaccaccacacatcgtggggaattggtctccaggtcacaggttctcttc 201
Qy 245 ttat 248
Db 202 atat 205

```

Search completed: August 7, 2002, 20:43:57  
 Job time: 4059 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2002, 18:55:53 : Search time 46.29 Seconds  
(without alignments)  
3162.618 Million cell updates/sec

Title: US-09-508-832-9  
Perfect score: 596  
Sequence: 1 atggcaagcaacctctga.....ctggtggaagaatgcatg 596

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2.6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2.6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2.6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2.6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2.6/ptodata/1/ina/PTUS.COMB.seq.\*  
6: /cgn2.6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	6.0	7218	1	US-08-232-463-14
2	33	5.5	1464	1	US-08-243-342-5
3	33	5.5	1464	1	US-08-477-407-5
4	33	5.5	1464	1	US-08-484-355-5
5	33	5.5	2913	1	US-08-243-542-7
6	33	5.5	2913	1	US-08-477-407-7
7	33	5.5	2913	1	US-08-484-355-7
8	33	5.5	2923	1	US-08-243-542-6
9	33	5.5	2923	1	US-08-477-407-6
10	33	5.5	2923	1	US-08-484-355-6
11	33	5.5	3183	1	US-08-243-542-8
12	33	5.5	3183	1	US-08-477-407-8
13	33	5.5	3183	1	US-08-484-355-8
14	32.2	5.4	3753	3	PCT-US95-02251-2
15	32.2	5.4	3759	3	US-08-479-7228-3
16	32.2	5.4	4314	1	US-08-199-780-2
17	32.2	5.4	4314	2	US-08-316-650-2
18	31.6	5.3	289	4	US-09-007-005-17
19	31.6	5.3	289	4	US-09-244-796-17
20	31.6	5.3	309	1	US-08-584-031-106
21	31.6	5.3	524	1	US-08-594-031-105
22	31.4	5.3	1509	1	US-08-216-276A-32
23	31.4	5.3	1622	1	US-08-216-276A-30
24	31.4	5.3	3230	1	US-08-216-276A-18
25	31.4	5.3	3230	1	US-08-219-2628-11
26	31.4	5.3	3230	3	US-09-031-655-11
27	31.4	5.3	3230	5	PCT-US91-03056-16

c 28 30.8 5.2 558 2 US-08-078-311-16  
c 29 30.8 5.2 558 2 US-08-460-402-16  
c 30 30.8 5.2 1907 1 US-08-184-327A-3  
c 31 30.8 5.2 1907 5 PCT-US95-00670-3  
c 32 30.8 5.2 2034 2 US-08-078-311-15  
c 33 30.8 5.2 2034 2 US-08-460-402-15  
c 34 30.6 5.1 1830 1 US-08-009-075-3  
c 35 30.2 5.1 2213 4 US-09-383-586-27  
c 36 30 5.0 576 4 US-09-385-982-23  
c 37 30 5.0 1941 3 US-09-082-737-1  
c 38 30 5.0 1987 2 US-08-990-379-1  
c 39 30 5.0 1993 2 US-08-990-379-2  
c 40 29.6 5.0 6226 1 US-08-542-363-1  
c 41 29.6 5.0 6226 4 US-09-100-089-1  
c 42 29.2 4.9 13987 2 US-08-804-227C-13  
c 43 29.2 4.9 44377 2 US-08-804-227C-7  
c 44 29.2 4.9 44377 2 US-08-804-198-1  
c 45 29.2 4.9 4403765 4 US-09-103-840A-2

## ALIGNMENTS

RESULT 1  
US-08-232-463-14  
: Sequence 14, Application US/08232463  
: Patent No. 5670367  
: GENERAL INFORMATION:  
: APPLICANT: DORNER, F.  
: APPLICANT: SCHEIFLINGER, F.  
: APPLICANT: FALKNER, F. G.  
: TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
: NUMBER OF SEQUENCES: 52  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Foley & Lardner  
: STREET: 1800 Diagonal Road, Suite 500  
: CITY: Alexandria  
: STATE: VA  
: COUNTRY: USA  
: ZIP: 22313-0299  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA: US/08/232,463  
: FILING DATE:  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US/07/935,313  
: FILING DATE:  
: APPLICATION NUMBER: EP 91 114 300.6  
: FILING DATE: 26-AUG-1991  
: ATTORNEY/AGENT INFORMATION:  
: NAME: BENT, Stephen A.  
: REGISTRATION NUMBER: 29,768  
: REFERENCE/DOCKET NUMBER: 30472/114 IMM  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (703)836-9300  
: TELEFAX: (703)683-4109  
: TELEX: 899149  
: INFORMATION FOR SEQ ID NO: 14:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 7218 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: IMMEDIATE SOURCE:  
: CLONE: pTZgpt-Fls  
: US-08-232-463-14

Sequence 16, Appl  
Sequence 16, Appl  
Sequence 3, Appl  
Sequence 3, Appl  
Sequence 15, Appl  
Sequence 15, Appl  
Sequence 3, Appl  
Sequence 27, Appl  
Sequence 23, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 2, Appl  
Sequence 1, Appl  
Sequence 13, Appl  
Sequence 7, Appl  
Sequence 1, Appl  
Sequence 2, Appl

```

Query Match      6.0%; Score 36; DB 1; Length 7218;
Best Local Similarity 3.8%; Pred. No. 0.12;
Matches 9; Conservative 138; Mismatches 93; Indels 0; Gaps 0;

Qy 162 ctccccccagcagccctcagggccgctggccacacccagccctgccccttttgc 221
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1212 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1271
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 222 tacagatccccgctttctatttatgagaagatccctcctgctgtctcagatccacg 281
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1272 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1331
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 282 tgggtattctcttttgacacagcagcaggagccacccatgagttgtgacaaatcaac 341
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1332 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1391
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 342 acaaaacccaaagctcctctccagcagcctcacaacactatctcagtgaaatggctccat 401
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1392 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1451
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
US-08-243-542-5
; Sequence 5, Application US/08243542
; Patent No. 5552526
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; TITLE OF INVENTION: MDC PROTEINS AND DNAs
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
; STREET: 2026 Rambling Road
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49008-1699
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT Compatible
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/243.542
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-136602
; FILING DATE: 14 MAY 1993
; APPLICATION NUMBER: JP 5-257455
; FILING DATE: 22 SEPTEMBER 1993
; APPLICATION NUMBER: JP 6-49904
; FILING DATE: 23 FEBRUARY 1994
; APPLICATION NUMBER: JP 6-73328
; FILING DATE: 12 APRIL 1994
; APPLICATION NUMBER: JP 6-84470
; FILING DATE: 22 APRIL 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Terryence F. Chapman
; REGISTRATION NUMBER: 32 549
; REFERENCE/DOCKET NUMBER: Furuya Case 1313
; TELEPHONE: (616) 381-1156
; TELEFAX: (616) 381-5465
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1464 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:

```

```

; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; LIBRARY: human fetal brain cDNA library
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1464
; US-08-243-542-5

Query Match      5.5%; Score 33; DB 1; Length 1464;
Best Local Similarity 55.8%; Pred. No. 0.5;
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 123 agttaacctgaagcaatcacgaggtgaaggagacagctgccccacagcagccctca 182
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 164 ATGGAACTTGACTTACATCGTGGAGCCCAAGAGGTGGCTGGACCTTGGGAGCCCTC 223
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 183 gggcccgctgccccacacccctgcccagcccttgccctttgtaccagatccccgc 235
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 224 AGGAGCCCTTCCACCCTCATTTACCGGACCCCTCTCCTCCAGATCCCTC 276
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
US-08-477-407-5
; Sequence 5, Application US/08477407
; Patent No. 5631351
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; TITLE OF INVENTION: MDC PROTEINS AND DNAs
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
; STREET: 2026 Rambling Road
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49008-1699
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT Compatible
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477.407
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/243.542
; FILING DATE: 13-MAY-1994
; APPLICATION NUMBER: JP 5-136602
; FILING DATE: 14 MAY 1993
; APPLICATION NUMBER: JP 5-257455
; FILING DATE: 22 SEPTEMBER 1993
; APPLICATION NUMBER: JP 6-49904
; FILING DATE: 23 FEBRUARY 1994
; APPLICATION NUMBER: JP 6-73328
; FILING DATE: 12 APRIL 1994
; APPLICATION NUMBER: JP 6-84470
; FILING DATE: 22 APRIL 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Terryence F. Chapman
; REGISTRATION NUMBER: 32 549
; REFERENCE/DOCKET NUMBER: Furuya Case 1313
; TELEPHONE: (616) 381-1156
; TELEFAX: (616) 381-5465
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1464 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```

TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE: human fetal brain cdna library  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE: human fetal brain cdna library  
LIBRARY: human fetal brain cdna library  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1464  
US-08-477-407-5

Query Match  
Best Local Similarity 55.8%; Pred. No. 0.5;  
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

OY 123 aggtatctcgaaggaacacacgaggtgaagggagacagctgcccacagcagcctca 182  
DB 164 ATGGGAACCTTACCTACATCGTGGAGCCCCCAAGAGGTGGCTGGAGCCCTC 223  
OY 183 gggccgcctggccccacacacgagcctggcccttttgcctaccagatccccgc 235  
DB 224 AGGAGACCCCTTCCCACTCATTTACCGGACCCCTCTCTCCAGATCCCTC 276

RESULT 4  
US-08-484-355-5  
Sequence 5, Application US/08484355  
Patent No. 5705341  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,355  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/243,542  
FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terryence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1464 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE: human fetal brain cdna library  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE: human fetal brain cdna library  
LIBRARY: human fetal brain cdna library  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1464  
US-08-484-355-5

Query Match  
Best Local Similarity 55.8%; Pred. No. 0.5;  
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

OY 123 aggtatctcgaaggaacacacgaggtgaagggagacagctgcccacagcagcctca 182  
DB 164 ATGGGAACCTTACCTACATCGTGGAGCCCCCAAGAGGTGGCTGGAGCCCTC 223  
OY 183 gggccgcctggccccacacacgagcctggcccttttgcctaccagatccccgc 235  
DB 224 AGGAGACCCCTTCCCACTCATTTACCGGACCCCTCTCTCCAGATCCCTC 276

RESULT 5  
US-08-243-542-7  
Sequence 7, Application US/08243542  
Patent No. 5552526  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/243,542  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terryence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 7:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2913 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; LIBRARY: human fetal brain cDNA library
; FEATURE:
; NAME/KEY: 5' UTR
; LOCATION: 1..27
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: 2038..2913
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..2037
;
US-08-243-542-7

Query Match          5.5%  Score 33; DB 1; Length 2913;
Best Local Similarity 55.8%; Pred. No. 0.73;
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 123 agttaatcctgaaggaatcagaggtgaagggacagctccccacggcagccctca 182
Db 212 ATGGGAACCTTGACTTACATCGTGGAGCCCAAGAGGTGGCTTGAGCCCTTGGGGAGCCCCCTC 271

Qy 183 gggcccgctggccacacgtccctggcccttttgcaccagatccccgc 235
Db 272 ACGGACCCCTTCCCACTCATTTACCGGACCCCTCTCTCCAGATCCCTC 324

RESULT 6
US-08-477-407-7
; Sequence 7, Application US/08477407
; Patent No. 5631351
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; TITLE OF INVENTION: MDC PROTEINS AND DNAS
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
; STREET: 2026 Rambling Road
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49008-1699
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,407
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/243,542
; FILING DATE: 13-MAY-1994
; APPLICATION NUMBER: JP 5-136602
; FILING DATE: 14 MAY 1993
; APPLICATION NUMBER: JP 5-257455
; FILING DATE: 22 SEPTEMBER 1993
; APPLICATION NUMBER: JP 6-49904
; FILING DATE: 23 FEBRUARY 1994
; APPLICATION NUMBER: JP 6-73328
; FILING DATE: 12 APRIL 1994
; APPLICATION NUMBER: JP 6-84470
```

```
; FILING DATE: 22 APRIL 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Terrivence F. Chapman
; REGISTRATION NUMBER: 32 549
; REFERENCE/DOCKET NUMBER: Furuya Case 1313
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (616) 381-1156
; TELEFAX: (616) 381-5465
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2913 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; LIBRARY: human fetal brain cDNA library
; FEATURE:
; NAME/KEY: 5' UTR
; LOCATION: 1..27
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: 2038..2913
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..2037
;
US-08-477-407-7

Query Match          5.5%  Score 33; DB 1; Length 2913;
Best Local Similarity 55.8%; Pred. No. 0.73;
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 123 agttaatcctgaaggaatcagaggtgaagggacagctccccacggcagccctca 182
Db 212 ATGGGAACCTTGACTTACATCGTGGAGCCCAAGAGGTGGCTTGAGCCCTTGGGGAGCCCCCTC 271

Qy 183 gggcccgctggccacacgtccctggcccttttgcaccagatccccgc 235
Db 272 ACGGACCCCTTCCCACTCATTTACCGGACCCCTCTCTCCAGATCCCTC 324

RESULT 7
US-08-484-355-7
; Sequence 7, Application US/08484355
; Patent No. 5705341
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; TITLE OF INVENTION: MDC PROTEINS AND DNAS
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
; STREET: 2026 Rambling Road
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49008-1699
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,355
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/243,542
; FILING DATE: 13-MAY-1994
```

```

: APPLICATION NUMBER: JP 5-136602
: FILING DATE: 14 MAY 1993
: APPLICATION NUMBER: JP 5-257455
: FILING DATE: 22 SEPTEMBER 1993
: APPLICATION NUMBER: JP 6-49904
: FILING DATE: 23 FEBRUARY 1994
: APPLICATION NUMBER: JP 6-73328
: FILING DATE: 12 APRIL 1994
: APPLICATION NUMBER: JP 6-84470
: FILING DATE: 22 APRIL 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Terryence F. Chapman
: REGISTRATION NUMBER: 32 549
: REFERENCE/DOCKET NUMBER: Furuya Case 1313
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (616) 381-1156
: TELEFAX: (616) 381-5465
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2913 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: IMMEDIATE SOURCE:
: LIBRARY: human fetal brain cDNA library
: FEATURE:
: NAME/KEY: 5' UTR
: LOCATION: 1..27
: FEATURE:
: NAME/KEY: 3' UTR
: LOCATION: 2038..2913
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 28..2037
:
: US-08-484-355-7

Query Match 5.5%; Score 33; DB 1; Length 2913;
Best Local Similarity 55.8%; Pred. No. 0.73;
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 123 agtaactcctgaaggcaatcacagggtgaaggagagagctgccccacggcagccctca 182
DB 212 ATGGGAACCTTGACTTACATCGTAGCCCCCAGAGGTGCTGGACCTTGGGAGCCCTC 271
QY 183 gggcccgctggccccaccctgccgcccctgttgcacagatccccgc 235
DB 272 AGGGACCCCTTCCCCACCTCATTTACCGGACCCCTCTCTCCACGATCCCTC 324

RESULT 8
US-08-243-542-6
: Sequence 6, Application US/08243542
: Patent No. 5552526
: GENERAL INFORMATION:
: APPLICANT: NAKAMURA, YUSUKE
: APPLICANT: : EMI, MITSURU
: TITLE OF INVENTION: MDC PROTEINS AND DNAs
: TITLE OF INVENTION: ENCODING THE SAME
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
: STREET: 2026 Rambling Road
: CITY: Kalamazoo
: STATE: Michigan
: COUNTRY: USA
: ZIP: 49008-1699
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
: COMPUTER: IBM PC/AT/Compatible

: APPLICATION NUMBER: JP 5-136602
: FILING DATE: 14 MAY 1993
: APPLICATION NUMBER: JP 5-257455
: FILING DATE: 22 SEPTEMBER 1993
: APPLICATION NUMBER: JP 6-49904
: FILING DATE: 23 FEBRUARY 1994
: APPLICATION NUMBER: JP 6-73328
: FILING DATE: 12 APRIL 1994
: APPLICATION NUMBER: JP 6-84470
: FILING DATE: 22 APRIL 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Terryence F. Chapman
: REGISTRATION NUMBER: 32 549
: REFERENCE/DOCKET NUMBER: Furuya Case 1313
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (616) 381-1156
: TELEFAX: (616) 381-5465
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2913 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: IMMEDIATE SOURCE:
: LIBRARY: human fetal brain cDNA library
: FEATURE:
: NAME/KEY: 5' UTR
: LOCATION: 1..27
: FEATURE:
: NAME/KEY: 3' UTR
: LOCATION: 2038..2913
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 28..2037
:
: US-08-484-355-7

Query Match 5.5%; Score 33; DB 1; Length 2913;
Best Local Similarity 55.8%; Pred. No. 0.73;
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 123 agtaactcctgaaggcaatcacagggtgaaggagagagctgccccacggcagccctca 182
DB 212 ATGGGAACCTTGACTTACATCGTAGCCCCCAGAGGTGCTGGACCTTGGGAGCCCTC 271
QY 183 gggcccgctggccccaccctgccgcccctgttgcacagatccccgc 235
DB 272 AGGGACCCCTTCCCCACCTCATTTACCGGACCCCTCTCTCCACGATCCCTC 324

RESULT 9
US-08-477-407-6
: Sequence 6, Application US/08477407
: Patent No. 5631351
: GENERAL INFORMATION:
: APPLICANT: NAKAMURA, YUSUKE
: APPLICANT: : EMI, MITSURU
: TITLE OF INVENTION: MDC PROTEINS AND DNAs
: TITLE OF INVENTION: ENCODING THE SAME
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
: STREET: 2026 Rambling Road
```



Oy 183 gggccgcgtggccacgtgcagccctggcccttttgcctaccagatccccgc 235  
 Db 272 AGGACCCCTTCCCACTTATTTACCGGACCCCTCTCTCCAGATCCCCCTC 324

RESULT 11

US-08-243-542-8  
 : Sequence 8, Application US/08243542  
 : Patent No. 5552526  
 : GENERAL INFORMATION:  
 : APPLICANT: NAKAMURA, YUSUKE  
 : TITLE OF INVENTION: MDC PROTEINS AND DNAs  
 : TITLE OF INVENTION: ENCODING THE SAME  
 : NUMBER OF SEQUENCES: 20  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
 : STREET: 2026 Rambling Road  
 : CITY: Kalamazoo  
 : STATE: Michigan  
 : COUNTRY: USA  
 : ZIP: 49008-1699  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
 : OPERATING SYSTEM: MS-DOS 5.0  
 : SOFTWARE: WordPerfect 5.0  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/243,542  
 : FILING DATE:  
 : CLASSIFICATION: 530  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: JP 5-136602  
 : FILING DATE: 14 MAY 1993  
 : APPLICATION NUMBER: JP 5-257455  
 : FILING DATE: 22 SEPTEMBER 1993  
 : APPLICATION NUMBER: JP 6-49904  
 : FILING DATE: 23 FEBRUARY 1994  
 : APPLICATION NUMBER: JP 6-73328  
 : FILING DATE: 12 APRIL 1994  
 : APPLICATION NUMBER: JP 6-84470  
 : FILING DATE: 22 APRIL 1994  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Terryence F. Chapman  
 : REGISTRATION NUMBER: 32 549  
 : REFERENCE/DOCKET NUMBER: Furuya Case 1313  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (616) 381-1156  
 : TELEFAX: (616) 381-5465  
 : INFORMATION FOR SEQ ID NO: 8:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 3183 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: double  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: cDNA to mRNA  
 : ORIGINAL SOURCE:  
 : ORGANISM: Homo sapiens  
 : IMMEDIATE SOURCE:  
 : LIBRARY: human fetal brain cDNA library  
 : FEATURE:  
 : NAME/KEY: 3' UTR  
 : LOCATION: 2308..3183  
 : FEATURE:  
 : NAME/KEY: CDS  
 : LOCATION: 1..2307  
 : US-08-243-542-8

Query Match 5.5%; Score 33; DB 1; Length 3183;  
 Best Local Similarity 55.8%; Pred. No. 0.77;  
 Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Oy 123 aggttaactctgaaggaatcaacgaggtgaagggagacagctgccccacacggcagccctca 182  
 Db 482 ATGGGAACCTTGACTTACATCGTGGAGAGCCCAAGAGGTGGCTGGACCTTGGGGAGCCCTC 541  
 Oy 183 gggccgcgtggccacgtgcagccctggcccttttgcctaccagatccccgc 235  
 Db 542 AGGACCCCTTCCCACTTATTTACCGGACCCCTCTCTCCAGATCCCCCTC 594

RESULT 12

US-08-477-407-8  
 : Sequence 8, Application US/08477407  
 : Patent No. 5631351  
 : GENERAL INFORMATION:  
 : APPLICANT: NAKAMURA, YUSUKE  
 : TITLE OF INVENTION: EMI, MITSURU  
 : TITLE OF INVENTION: MDC PROTEINS AND DNAs  
 : TITLE OF INVENTION: ENCODING THE SAME  
 : NUMBER OF SEQUENCES: 20  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
 : STREET: 2026 Rambling Road  
 : CITY: Kalamazoo  
 : STATE: Michigan  
 : COUNTRY: USA  
 : ZIP: 49008-1699  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
 : OPERATING SYSTEM: IBM PC/XT/AT Compatible  
 : SOFTWARE: WordPerfect 5.0  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/477,407  
 : FILING DATE: 07-JUN-1995  
 : CLASSIFICATION: 530  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/243,542  
 : FILING DATE: 13-MAY-1994  
 : APPLICATION NUMBER: JP 5-136602  
 : FILING DATE: 14 MAY 1993  
 : APPLICATION NUMBER: JP 5-257455  
 : FILING DATE: 22 SEPTEMBER 1993  
 : APPLICATION NUMBER: JP 6-49904  
 : FILING DATE: 23 FEBRUARY 1994  
 : APPLICATION NUMBER: JP 6-73328  
 : FILING DATE: 12 APRIL 1994  
 : APPLICATION NUMBER: JP 6-84470  
 : FILING DATE: 22 APRIL 1994  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Terryence F. Chapman  
 : REGISTRATION NUMBER: 32 549  
 : REFERENCE/DOCKET NUMBER: Furuya Case 1313  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (616) 381-1156  
 : TELEFAX: (616) 381-5465  
 : INFORMATION FOR SEQ ID NO: 8:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 3183 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: double  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: cDNA to mRNA  
 : ORIGINAL SOURCE:  
 : ORGANISM: Homo sapiens  
 : IMMEDIATE SOURCE:  
 : LIBRARY: human fetal brain cDNA library  
 : FEATURE:  
 : NAME/KEY: 3' UTR  
 : LOCATION: 2308..3183  
 : FEATURE:  
 : NAME/KEY: CDS  
 : LOCATION: 1..2307  
 : US-08-477-407-8

```
Query Match          5.5%  Score 33; DB 1; Length 3183;
Best Local Similarity 55.8%; Pred. No. 0.77; 50; Indels 0; Gaps 0;
Matches 63; Conservative 0; Mismatches 0; Gaps 0;

QY 123 agttaactgaaggaacacacgaggtgaaggagacagctgccccacggcagccctca 182
DB 482 ATGGAACTTACTTACATCTGAGCCCAAGAGGTGCTGGACCTTGGGGAGCCCTC 541
QY 183 gggcccgctggccccacacgacccctgtgcccccttttctaccagatccccgc 235
DB 542 AGGACCCCTTCCCACTTATTCAGCGACCCCTCTCTCTCCACATCCCTC 594

RESULT 13
US-08-484-355-8
: Sequence 8, Application US/08484355
: Patent No. 5705341
: GENERAL INFORMATION:
: APPLICANT: NAKAMURA, YUSUKE
: TITLE OF INVENTION: MDC PROTEINS AND DNAs
: TITLE OF INVENTION: ENCODING THE SAME
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
: STREET: 2026 Rambling Road
: CITY: Kalamazoo
: STATE: Michigan
: COUNTRY: USA
: ZIP: 49008-1699
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
: COMPUTER: IBM PC/XT/AT Compatible
: OPERATING SYSTEM: MS-DOS 5.0
: SOFTWARE: WordPerfect 5.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/484,355
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/243,542
: FILING DATE: 13-MAY-1994
: APPLICATION NUMBER: JP 5-136602
: FILING DATE: 14 MAY 1993
: APPLICATION NUMBER: JP 5-257455
: FILING DATE: 22 SEPTEMBER 1993
: APPLICATION NUMBER: JP 6-49904
: FILING DATE: 23 FEBRUARY 1994
: APPLICATION NUMBER: JP 6-73328
: FILING DATE: 12 APRIL 1994
: APPLICATION NUMBER: JP 6-84470
: FILING DATE: 22 APRIL 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Terrence F. Chapman
: REGISTRATION NUMBER: 32 549
: REFERENCE/DOCKET NUMBER: Furuya Case 1313
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (616) 381-1156
: TELEFAX: (616) 381-5465
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3183 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: IMMEDIATE SOURCE:
: LIBRARY: human fetal brain cDNA library
: FEATURE:
```

```
: NAME/KEY: 3' UTR
: LOCATION: 2308...3183
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..2307
: US-08-484-355-8

Query Match          5.5%  Score 33; DB 1; Length 3183;
Best Local Similarity 55.8%; Pred. No. 0.77;
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 123 agttaactgaaggaacacacgaggtgaaggagacagctgccccacggcagccctca 182
DB 482 ATGGAACTTACTTACATCTGAGCCCAAGAGGTGCTGGACCTTGGGGAGCCCTC 541
QY 183 gggcccgctggccccacacgacccctgtgcccccttttctaccagatccccgc 235
DB 542 AGGACCCCTTCCCACTTATTCAGCGACCCCTCTCTCTCCACATCCCTC 594

RESULT 14
PCT-US95-02251-2/C
: Sequence 2, Application PC/TUS9502251
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
: NUMBER OF SEQUENCES: 18
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Arnold, White & Durkee
: STREET: P.O. Box 4433
: CITY: Houston
: STATE: Texas
: COUNTRY: United States of America
: ZIP: 77210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
: SOFTWARE: Patent In Release #1.0, Version
: SOFTWARE: #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/02251
: FILING DATE: CONCURRENTLY HEREMITH
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/316,650
: FILING DATE: 30-SEP-1994
: CLASSIFICATION:
: APPLICATION NUMBER: US 08/199,780
: FILING DATE: 18-FEB-1994
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Parker, David L.
: REGISTRATION NUMBER: 32,165
: REFERENCE/DOCKET NUMBER: UMIC009P--
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (512) 418-3000
: TELEFAX: (713) 789-2679
: TELEFAX: 79-0924
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3753 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..3753
: PCT-US95-02251-2
```



Query Match 5.4%; Score 32.2; DB 5; Length 3753;  
 Best Local Similarity 52.6%; Pred. No. 1.5;  
 Matches 70; Conservative 0; Mismatches 63; Indels 0; Gaps 0;  
 OY 77 cccagctcagacctg99gcccctacctccctacagacagagccacaaagtgatctctgaag 136  
 DB 200 CCTGTGTGACAGCTGTCCCGACACTGGCCCTTCAGACAGGTCCGCTTCAGATCACAGGC 141  
 OY 137 gcaatcacgaggtgaaggggagcagctgccccacgagcagcctcagggcccgctggccc 196  
 DB 140 GCAAAGACCACCTTGAAGCGTTGGGCCCGCAGCGCCGCCGCCCTGTGCCCCGCTGCC 81  
 OY 197 caactgcccagccc 209  
 DB 80 GGCGGCCCCACCC 68

RESULT 15  
 US-08-479-722B-3/c  
 ; Sequence 3, Application US/08479722B  
 ; Patent No. 6074840  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bonadio, Jeffrey  
 ; APPLICANT: Yin, Wushan  
 ; TITLE OF INVENTION: LATENT TGF( BINDING PROTEIN (LTBP)  
 ; TITLE OF INVENTION: GENES, COMPOSITIONS AND METHODS  
 ; NUMBER OF SEQUENCES: 13  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Williams, Morgan & Amerson  
 ; STREET: 7676 Hillmont, Suite 250  
 ; CITY: Houston  
 ; STATE: Texas  
 ; COUNTRY: USA  
 ; ZIP: 77040  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/479,722B  
 ; FILING DATE: 07-JUN-1995  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US PCT/US95/02251  
 ; FILING DATE: 21-FEB-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/316,650  
 ; FILING DATE: 30-SEP-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/199,780  
 ; FILING DATE: 18-FEB-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Fussey, Shelley P. M.  
 ; REGISTRATION NUMBER: 39,458  
 ; REFERENCE/DOCKET NUMBER: 4100.000500/FUS  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (713) 934-7000  
 ; TELEFAX: (713) 934-7011  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3759 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..3759  
 ; US-08-479-722B-3

Query Match 5.4%; Score 32.2; DB 3; Length 3759;

Best Local Similarity 52.6%; Pred. No. 1.5;  
 Matches 70; Conservative 0; Mismatches 63; Indels 0; Gaps 0;  
 OY 77 cccagctcagacctg99gcccctacctccctacagacagagccacaaagtgatctctgaag 136  
 DB 203 CCTGTGTGACAGCTGTCCCGACACTGGCCCTTCAGACAGGTCCGCTTCAGATCACAGGC 144  
 OY 137 gcaatcacgaggtgaaggggagcagctgccccacgagcagcctcagggcccgctggccc 196  
 DB 143 GCAAAGACCACCTTGAAGCGTTGGGCCCGCAGCGCCGCCGCCCTGTGCCCCGCTGCC 84  
 OY 197 caactgcccagccc 209  
 DB 83 GGCGGCCCCACCC 71

Search completed: August 7, 2002, 20:40:18  
 Job time: 6265 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2002, 18:47:18 : Search time 1730.67 seconds  
(without alignments)  
4648.018 million cell updates/sec

Title: US-09-508-832-9  
Perfect score: 596  
Sequence: 1 atggcaagaacactctctga.....ctgggtggagaatgcattg 596

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- EST:\*
- 1: em\_estba.\*
  - 2: em\_estchun.\*
  - 3: em\_estin.\*
  - 4: em\_estnu.\*
  - 5: em\_estov.\*
  - 6: em\_estpl.\*
  - 7: em\_estro.\*
  - 8: em\_hic.\*
  - 9: gb\_est1.\*
  - 10: gb\_est2.\*
  - 11: gb\_hic.\*
  - 12: gb\_gss.\*
  - 13: gb\_gss\_hum.\*
  - 14: em\_gss\_inv.\*
  - 15: em\_gss\_pln.\*
  - 16: em\_gss\_vrt.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	430	72.1	1206	11	AK011490	Mus muscu
2	379.4	63.7	935	10	BG921698	602825518
3	359.6	60.3	700	10	B1905766	603064506
4	330	55.4	580	12	A2706148	RPCI-23-2
5	329.4	55.3	645	9	B8651225	B8651225
6	328.4	55.1	686	9	B8577135	B8577135
7	315.4	52.9	452	10	BF021882	BF021882
8	250.8	42.1	668	10	BG173095	BG173095
9	186.6	33.0	765	10	B1454637	B1454637
10	185.6	31.1	480	9	B8556627	B8556627
11	185.2	31.1	684	9	B8653788	B8653788
12	124	20.8	492	9	A1971169	wr24hl2.x
13	112.8	18.9	620	9	B8631162	B8631162
14	112.8	18.9	912	10	B1555336	603236095
15	99	16.6	135	10	B1497208	B1497208
16	98	16.4	501	9	AA629308	zu84g06.s
17	97	16.3	501	9	AA629050	zu84a06.s

18	94.8	15.9	157	9	AW748960	AW748960
19	94.6	15.9	584	9	AW629314	AW629314
20	91.8	15.4	210	10	BF172831	PCL5805.M
21	84.2	14.1	389	10	BF319454	uy59b09.x
22	77.4	13.0	97	9	A1970428	wr10d03.x
23	68.4	11.5	289	10	BG990772	BG990772
24	53.4	9.0	537	9	AF209718	AF209718
25	41.4	6.9	1101	12	CNS00370	CNS00370
26	38.6	6.5	342	9	AW531398	UI-R-BSO-
27	38.6	6.5	964	12	CNS003MG	AL065254
28	38.4	6.4	925	12	CNS0091P	Drosophila
29	37.8	6.3	450	10	B2515545	AL051013
30	37.4	6.3	504	9	AV621963	Drosophila
31	37.4	6.3	639	10	BM000373	WHE0614.D
32	37.4	6.3	651	10	B1718793	1031031D1
33	37.4	6.3	662	10	B1724305	1031069H1
34	37.4	6.3	685	10	B1717797	1031022A1
35	37.4	6.3	1545	10	BM467095	BM467095
36	37.2	6.2	512	9	BB867390	BB867390
37	37	6.2	643	10	BE295393	601175911
38	37	6.2	717	10	B1261466	602853519
39	37	6.2	894	10	B1114718	602861747
40	36.8	6.2	298	10	B6275727	601211880
41	36.8	6.2	773	10	B6421208	602451695
42	36.4	6.1	587	9	AJ003454	AJ003454
43	36.4	6.1	925	12	CNS0091P	AL053013
44	36.2	6.1	284	9	AI473314	Drosophila
45	36.2	6.1	414	10	BF850601	th83d09.x
						BF850601

## ALIGNMENTS

### RESULT

AK011490

LOCUS

DEFINITION

AK011490

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

AK011490 1206 bp mRNA llinear HTC 19-JAN-2002  
Mus musculus 10 days embryo whole body cDNA, RIKEN full-length  
enriched library, clone:2610020M23:BCL2-like 11 (apoptosis  
facilitator), full insert sequence.  
AK011490 1 GI:12847647  
HTC: CAP trapper.  
Mus musculus (strain:C57BL/6J) 10 days embryo cDNA to mRNA,  
clone\_11B:RIKEN full-length enriched mouse CDNA library  
clone:2610020M23.  
Mus musculus  
Eukaryote: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:  
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.  
1 (sites)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303:19-44 (1999)  
99279253  
10349636  
2 (sites)  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159  
3 (sites)  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Hatada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)



/db\_xref="taxon:10090"  
 /clone="IMAGE:4954300"  
 /clone\_lib="NCI\_GCAP\_Mam6"  
 /sex="female, virgin"  
 /tissue\_type="infiltrating ductal carcinoma"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Jeffrey Green, M.D., NIH"  
 BASE COUNT 203 a 283 c 276 g 173 t  
 ORIGIN

Query Match 63.7%; Score 379.4; DB 10; Length 935;  
 Best Local Similarity 85.7%; Pred. No. 2.9e-97;  
 Matches 510; Conservative 0; Mismatches 61; Indels 24; Gaps 7;  
 QY 1 atggcaagcaacctctctgatgaagttctgtgagtgtagcaggaagtagacaattgcag 60  
 Db 208 ATGGCAAGCAACCTTCTGATGTAAGTCTGTGAGTGTGACAGAGAAGGTGCACAATTGCAG 267  
 QY 61 cctcgagagagctcccgagctgacctggggccctacctccctcacagagagcca 120  
 Db 268 CTTCTGACAGAGGCTCCACAGCTGAGGCTGGGGCCCTACTCTCTACAGACAGAACCG 327  
 QY 121 caaggttaacctgaaggcaatcacggaggtgaaggggagacagctgccccacacgca-gccc 179  
 Db 328 CAAGGTAAATCCGACGCG-----GAAGGGGACCGCTGCCGCCACGCGCATGCC 375  
 QY 180 tcaggggccgctggccccacacctgcccagccctggccc-ttttgcaccagatcccgcctt 238  
 Db 376 TCAGGGCCCGCTGGCCCCACGCGCCCTGGCCCTTTTGCTACCGAGATCCCACTTT 435  
 QY 239 tcatttttagaagaatctccctgctgctgctgctgctgctgctgctgctgctgctgctgctg 298  
 Db 436 TCATCTTTTGAGAAATCTCTCTGCTGCTCCGGCTCTCCAGTGGGTATTTCTCTTTTG 495  
 QY 299 acacagacagagagccagcaccatgagttgtgacaaatcaacacaaaccccaagtcctc 358  
 Db 496 ACACAGACAGAGAGCCCGGACCCATGAGTGTGTGACAACTCAACACAA-CCCAAGTCCCTC 554  
 QY 359 cttgcccagggcttcaaccactctcagtgcaatgagcttccatgagcagagctc-----g 412  
 Db 555 CTTGCCAGGCTTCAACCACTATCTCAGTGCATGCTTCCCTACGACACTCTCAGGAGG 614  
 QY 413 aacctgcagatacgccccagagatatgagtcgccccagaggttgcgggtatcgagagcg 472  
 Db 615 AACCTGAAGATCTGCCGCCGGAGATACGGATTGCACAGGAGCTCGCGCGGATCGGAGAG 674  
 QY 473 agtttaacgcttactatgcaagaggggtatttttgaataattacacagcagccgagagcc 532  
 Db 675 AGTTCAAGAAACTTACACAGAGAGGGGTGTTTGCAAAATGATTACCGGAGCGTGAAGA-C 733  
 QY 533 eccacagaatgggtattcttcaag--ctgttaagttacatatgtccgcctgggtgag 585  
 Db 734 ACCCTCAATGGTATTATTACAAACGTGTGTACGCTTATCTTCGCTGTGTATGG 788

RESULT 3  
 BI906766  
 LOCUS 700 bp mRNA linear EST 16-OCT-2001  
 DEFINITION 603064506F1 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5213713 5',  
 mRNA sequence.  
 ACCESSION BI906766  
 VERSION BI906766.1 GI:16169524  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 700)

NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation by: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM11536 row: 1 column: 02  
 High quality sequence stop: 696.  
 FEATURES  
 source  
 1. 700  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5213713"  
 /clone\_lib="NIH\_MGC\_118"  
 /tissue\_type="Leukocyte"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV  
 (destroyed); RNA source leukocytes from anonymous pool of  
 non-activated adult donors. Library is oligo-dT primed  
 and directionally cloned (EcoRV site is destroyed upon  
 cloning). Average insert size 1.7 kb, insert size range  
 1.2-3.3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 027. Note:  
 this is a NIH\_MGC Library."  
 BASE COUNT 148 a 233 c 175 g 144 t  
 ORIGIN

Query Match 60.3%; Score 359.6; DB 10; Length 700;  
 Best Local Similarity 96.2%; Pred. No. 1.1e-91;  
 Matches 433; Conservative 0; Mismatches 9; Indels 8; Gaps 6;  
 QY 1 atggcaagcaacctctctgatgaagtt-ctgagtgtagcaggaagtagacaattgcga 59  
 Db 257 ATGGCAAGCAACCTTCTGATGTAAGTCTGTGAGTGTGACGAGAAGGTAGACAATTGCA 316  
 QY 60 gctcgagagagggctcccgagctcagacctggggccctacctccctcacagacagagcc 119  
 Db 317 GCCTCGGAGAGGGCTCCCGAGCTCAGACCTGGGGCCCCCTACTCTCCCTACAGACAGAGCC 376  
 QY 120 acaaggttaacctgaaggcaatcacggaggtgaagggggagagctgccccccagcagccc 179  
 Db 377 ACAAGGTAACTCTGAAGGCAATCACGGAGGTGAAGGGGACAGCTGCCCCACGAGCCTC 436  
 QY 180 tcaggggccgctggccccacacctgccagcctggccttttgcctaccagatcccgccttt 239  
 Db 437 ---AGGGCCCGTGGCCCCACCTCGCCAGCCCTGG-CTTTTGTACCAATTCGCCGCTTTT 492  
 QY 240 catcttttagaagaagatctccctcgtgctcgtatcctccagtggtggtattctcttttga 299  
 Db 493 CACTCTTATGAGAGATCTCTCCCTCGTGTCTCATCTCTCCAGTGGGTGTA-TTCTCTTTTGA 551  
 QY 300 cacagacagggggcccgagcaccatgagttgtgacaaatcaacacaaaccccaagtcctcc 359  
 Db 552 CACAGACAGGAGCCGACGCCCATGAGTTGTGACAAATCAACACAAACCCCAAGTCTCTCC 611  
 QY 360 ttgcagagccttcaaccactatctcagtcacatgcttcctcagagaggtgacctgc 419  
 Db 612 TTGCCAGGCCCTTCAA-CACTATCTCAGTGCATGGGCTTCATGAGGAGGCTGAACCTGG 670  
 QY 420 agatatgcgcccagagatatgatgcgcca 449  
 Db 671 AGATATGCC-CCAGAGATATGATGCCCA 699  
 RESULT 4



```

/clone-lib="RIKEN full-length enriched, 0 day neonate
cerebellum"
/tissue_type="cerebellum"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/notes-Site:1: Sali; Site:2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGAGATCCAGAGCTCTTTTNN 3']. cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 479.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGATCTCGAGTTAATTAATCCCGCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I.

BASE COUNT      130 a 229 c 156 g 130 t
ORIGIN

Query Match      55.3% Score 329.4; DB 9; Length 645;
Best Local Similarity 89.7%; Pred. No. 4.5e-83;
Matches 394; Conservative 0; Mismatches 26; Indels 19; Gaps 3;

QY 1 atggcaagaacacctctgatgttgatgtgacgagagagtgagacaaattgcag 60
Db 220 ATGGCCAGAACCTTCTGATGTAGTCTGAGTGTGACAGAGAGGTGGACAATTGCAG 279

QY 61 cctgcgagagggctcccgagctcagacctgggcccctacctccctacagacagagcc 120
Db 280 CTCTCTGAGAGGCTCCCGAGCTCAGGCTGGGGCCCTTACTCTCTACAGACAGAACGG 339

QY 121 caaggttaactgaaggacatcacaggaggtgaaggagagagctccccccagcagccct 180
Db 340 CAAGGTAACTCCGACGGC-----GAAGGGGAGACCGCTGCCCCACAGCAGCCCT 387

QY 181 caggggccgcgtggccacacgtccagccctggcccttttctacagatccccgcttttc 240
Db 388 CAGGGCCCGCTGGCCCGCCAGCGGCGCTGGGCCCTTTTGTCTACAGATCCCCACTTTC 447

QY 241 atctttgagagagatctctctgtgtctcgatctccagtgagggtattctctttgac 300
Db 448 ATCTTTGTGAGAGATCTTCTGTCTGCCGGCTCCCTCCAGTGGGTATTTCTCTTTTGAC 507

QY 301 acagacaggagccagcaccatgattgtgacaaatcaacacaaaccccaagctccctct 360
Db 508 ACAGACAGGAGCGCGCCACCATGATTGTGACAAAGTCAACACAAACCCCAAGTCTCCT 567

QY 361 tgcagagcctcaaccactatctcagtgcaatggcttccatgagcaggtc-----gaa 414
Db 568 TGCAGAGCCCTTC-ACCATATCTCAGTCCATGGCTCCATACACACATGCTCAGGAGGAA 626

QY 415 cctcagagatgcgcccag 433
Db 627 CCTGAAGATCTGCGCCCGG 645

RESULT 6
LOCUS      BB577135      686 bp      mRNA      linear      EST 26-OCT-2001
DEFINITION BB577135 RIKEN full-length enriched, 13 days embryo male testis MUS
ACCESSION  BB577135
VERSION    BB577135.2 GI:16449433
KEYWORDS   EST.
SOURCE     house mouse.

```

## ORGANISM

REFERENCE  
AUTHORS

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (Bases 1 to 686)  
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,  
Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,  
M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,  
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,  
D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,  
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,  
Muramatsu,M. and Hayashizaki,Y.  
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

TITLE  
JOURNAL  
COMMENT

Unpublished (2001)  
On Nov 30, 2000 this sequence version replaced gi:11473679.  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Sucho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-resgsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/  
Carninci,P., Shibata,Y., Hayatsu,M., Sugahara,Y., Shibata,K., Itoh,  
M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
Wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,  
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,  
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and  
Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,  
Y. and Hayashizaki,Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Konno,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,  
K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and  
Hayashizaki,Y.  
Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Please visit our web site (http://genome.gsc.riken.go.jp/) for  
further details.  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN,  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

## FEATURES

## source

Location/Qualifiers  
1..686  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_lib="RIKEN full-length enriched, 13 days embryo  
male testis"  
/sex="male"  
/tissue\_type="testis"  
/dev\_stage="13 days embryo"  
/lab\_host="DH10B"  
/note="Site:1: Sali; Site:2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN, Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5',  
GAGAGAGAGATCCAGAGCTCTTTTNN 3']. cDNA was  
prepared by using trehalose thermo-activated reverse







Oy 596 q 596  
Db 594 g 594

## RESULT 10

BB856627  
DEFINITION  
BB856627 RIKEN full-length enriched, B16 F10Y cells mRNA linear EST 26-NOV-2001  
CDNA clone G370035M16 5', mRNA sequence.

## ACCESSION

BB856627

## VERSION

BB856627.1

## KEYWORDS

EST.

## SOURCE

house mouse.

## ORGANISM

Mus musculus

## REFERENCE

1 (bases 1 to 480)

## AUTHORS

Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Ishii, K., Kawauchi, T., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numagami, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akanishi, S., Tanaka, T., Tomaru, A., Toya, T., Watanabe, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

## TITLE

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

## JOURNAL

unpublished (2001)

## COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Carninci, P., Shibata, K., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh

, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

Watanabe, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura

, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and

Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara

, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Please visit our web site (http://genome.gsc.riken.go.jp) for

further details.

## FEATURES

Location/Qualifiers

1. 480

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="G370035M16"

/clone\_lib="RIKEN full-length enriched, B16 F10Y cells"

/cell\_type="B16 F10Y cells"

88 a 175 c 121 g 96 t

## BASE COUNT

ORIGIN

Query Match 31.1% Score 185.6; DB 9: Length 480;  
Best Local Similarity 87.5%; Pred. NO. 4e-42;

Matches 231; Conservative 0; Mismatches 19; Indels 14; Gaps 2;

Oy 1 atggcaagcaacccctctgattgaagttctgagtgtagccgagaaggtagacaattgag 60  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 230 ATGCCAAGCAACCTTCGTGATGTAGTTCGTAGTGTGACAGAGAGTGGCAATGGAG 289  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Oy 61 cctg-cyggagggccctccccagctcagacctggggggccctcctccctacagacagagcc 119  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 290 CTTGCTGTAGAGGCGCTCCAGCTCAGGCGCTGGGGCGCCCTACCTCCCTACAGACAGA --- 346  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Oy 120 acaaggaatactcctgaagcaatcacggaggtggaagggagcagctgccccacggcagcc 179  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 347 -----ACGCAAGGTATATCCGACGGCGAAGGGGACCGCTGCCCCACGCGAGCCC 396  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Oy 180 tcaggggcccgctggcccaactgcccagccctggccctttgtcttaacagatccccgctttt 239  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 397 TCAGGGCCCGCTGGGCGCCACCGGCGAGCCCTGGCCCTTTTCTTACCAGATGCCCACTTTT 456  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Oy 240 catctttatgagaagatctctcctc 263  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 457 CATCTTTGTGAGAAGATCTTCTCT 480  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

## RESULT 11

BB853788

## LOCUS

BB853788

## DEFINITION

musculus cDNA clone C730026B05 5', mRNA sequence.

## ACCESSION

BB853788

## VERSION

BB853788.1

## KEYWORDS

EST.

## SOURCE

house mouse.

## ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 684)

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.

, Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda

, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,

Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki

, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,

Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,

Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

## TITLE

unpublished (2001)

## JOURNAL

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Carninci, P., Shibata, K., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh

, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

Watanabe, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura

, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and

Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara

, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Please visit our web site (http://genome.gsc.riken.go.jp) for

further details.

e mouse tissues.





Qy 61 cctgcgagcgctcccgctcagacgtggccctaccctccctacagacagcca 120  
 Db 321 cctgcctgagagcgctcccgctcagcgtggccctccctaccctccctacagacagcca 380  
 Qy 171 caag 124  
 Db 381 CAAG 384

Search completed: August 7, 2002, 20:39:02  
 Job time: 6704 sec

RESULT 15  
 BI497208  
 LOCUS  
 DEFINITION df133e06.y1 Morton Fetal Cochlea Homo sapiens cDNA clone  
 IMAGE:2537938 5', mRNA sequence.  
 ACCESSION BI497208  
 VERSION BI497208.1 GI:15336552  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 135)  
 AUTHORS Robertson, N.G., Khetarpal, U., Gutierrez-Espeleta, G.A., Bieber, F.R.,  
 and Morton, C.C.  
 TITLE Isolation of novel and known genes from a human fetal cochlear cDNA  
 library using subtractive hybridization and differential screening  
 JOURNAL Genomics 23, 42-50 (1994)  
 MEDLINE 95130111  
 COMMENT Contact: Morton, C. C.  
 Departments of Pathology and Obstetrics, Gynecology and  
 Reproductive Biology  
 Brigham and Women's Hospital  
 75 Francis Street, Harvard Medical School, Boston, MA 02115, USA  
 Tel: 617 732 7980  
 Fax: 617 738 6996  
 Email: cmorton@rics.bwh.harvard.edu  
 DNA sequencing and analyses were performed by National Institutes  
 of Health Intramural Sequencing Center (NISC; see  
 http://www.nisc.nih.gov).  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Plate: LLAM6321 row: I column: 11  
 Seq primer: M13RP1 reverse primer (AB1).  
 FEATURES  
 source  
 1..135  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2537938"  
 /clone\_lib="Morton Fetal Cochlea"  
 /tissue\_type="cochlea"  
 /dev\_stage="16-22 week fetus"  
 /lab\_host="SOLR cells (kanamycin resistant)"  
 /note="Organ: ear; Vector: pBluescript SK-; Site: 1: EcoRI;  
 Site: 2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned  
 unidirectionally. Primer: Oligo dT. Fetal cochlea, normal.  
 37% of inserts <0.5 kb. 568 0.5-1.0 kb. 78 >1 kb. Uni-ZAP.  
 XR Vector. Library constructed by N. Robertson, C. Morton.  
 -5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor  
 sequence: 5' CTCGAGTTTCTTTTCTTTT j."

BASE COUNT 32 a 29 c 35 g 39 t  
 ORIGIN  
 Query Match 16.6% Score 99; DB 10; Length 135;  
 Best Local Similarity 100.0%; Pred. No. 1.le-17;  
 Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 498 ggtatttttgaataataacagcagcgagacacccacacgaatggtattcttaccact 557  
 Db 17 GGTATTTTGAATAATACAGCAGCGAGACACCCACGAAATGTTATCTTACGACT 76  
 Qy 558 gttacgttacattgctcgcctggtgtggaagtgcattg 596  
 ||||||||||||||||||||||||||||||||||||||||

OM of: US-09-508-832-10 to: GenEmbl.\* out\_format : pfs

Date: Aug 7, 2002 10:35 PM

About: Results were produced by the GenCore software, version 4.5.  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=Cgn2\_1/USPTO\_spool/US09508832/runat\_06082002\_103611\_2778/app\_query.fasta\_1.259  
-DB=GenEmbl -OEMT=Eastap -SUFFIX=xe -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -OGAPOP=4.500  
-OGAPEXT=0.050 -XGAPOP=0.000 -XGAPEXT=0.500 -FGAPOP=6.000  
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-DELEXT=7.000 -START=1 -NATRX=BIOSUM62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0  
-ALIGN=15 -MODE=LOCAL -OUTEMT=pfs -NORM=ext -HEAFSIZE=500  
-MINLEN=0 -MAXLEN=200000000 -USER=US09508832\_ECGL1\_3954  
-NCPU=6 -ICPU=3 -LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30  
-NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-508-832-10  
Query length: 198  
Database: GenEmbl.\*  
Database sequences: 1797656  
Database length: 187333701  
Search time (sec): 1829.680000

score\_list:

Sequence	Strd	Orig	zscore	EScore	Len	Document
gb_pat:AX031287	+	1065.00	1039.47	1.6e-49	596	AX031287 Sequence 9 from Patent
gb_un:AX031313	+	1065.00	1039.47	1.6e-49	596	AX031313 Sequence 9 from Patent
gb_pat:AF071197	+	1065.00	1039.46	1.6e-49	597	AF071197 Homo sapiens BimL mRNA
gb_un:AF071195	+	1040.50	1024.15	1.1e-48	630	AF071195 Homo sapiens mRNA for
gb_pat:AX031283	+	1040.50	985.61	1.6e-46	722	AX031283 Sequence 5 from Patent
gb_un:AX031309	+	908.00	888.51	1.0e-41	590	AX031309 Sequence 5 from Patent
gb_un:AF032459	+	908.00	888.49	1.0e-41	591	AF032459 Mus musculus BimL mRNA
gb_ro:BC019556	+	908.00	883.63	7.4e-41	1137	BC019556 Mus musculus, BCL2-1
gb_ro:AF065433	+	893.00	874.06	2.5e-40	591	AF065433 Rattus norvegicus Bcl-1
gb_pat:AF071198	+	873.50	822.02	2.0e-37	493	AF071198 Homo sapiens mRNA for
gb_pat:AC096670	+	713.00	658.32	2.6e-28	181497	AC096670 Homo sapiens chromo
gb_un:AX031285	+	696.00	687.16	6.5e-30	416	AX031285 Sequence 7 from Patent
gb_pat:AX031311	+	696.00	687.16	6.5e-30	416	AX031311 Sequence 7 from Patent
gb_pat:AF032458	+	641.50	632.76	7.0e-27	542	AF032458 Homo sapiens BimL mRNA
gb_un:AF071196	+	589.00	584.11	3.6e-24	422	AF071196 Homo sapiens mRNA for
gb_pat:AX031281	+	589.00	584.11	3.6e-24	422	AX031281 Sequence 3 from Patent
gb_un:AF032460	+	589.00	584.11	3.6e-24	423	AF032460 Mus musculus BimL mRNA
gb_ro:AF032460	+	574.00	569.67	2.3e-23	423	AF032460 Rattus norvegicus Bcl-1
gb_pat:AF071199	+	535.00	531.80	2.9e-21	443	AF071199 Homo sapiens mRNA for
gb_un:AX031279	+	406.00	409.85	1.8e-14	332	AX031279 Sequence 1 from Patent
gb_un:AX031305	+	406.00	409.85	1.8e-14	332	AX031305 Sequence 1 from Patent
gb_ro:AF032461	+	388.00	392.51	1.7e-13	333	AF032461 Mus musculus BimL mRNA
gb_ro:AF065432	+	388.00	388.93	2.7e-13	282	AF065432 Rattus norvegicus Bcl-1
gb_ro:AF065431	+	285.50	306.50	1.0e-08	223	AF065431 Rattus norvegicus Bcl-1
gb_pat:AF071200	+	201.00	199.10	0.0100	2050	AF071200 Homo sapiens mRNA for
em_hcg_hum:AC046192	-	192.00	158.53	1.82	150035	AC046192 Homo sapiens chromo
em_hcg_hum:AC046192	-	192.00	157.75	2.01	166638	AC046192 Homo sapiens chromo
em_hcg_hum:AC046192	-	170.00	137.36	27.44	150035	AC046192 Homo sapiens chromo
gb_hcg:AC104222	-	128.00	103.79	2.0e+03	59827	AC104222 Mus musculus clone 4
gb_pat:AC004907	-	125.50	114.27	530.19	10570	AC004907 Pseudomonas aerugin
gb_pat:AC004971	-	122.50	93.13	8.0e+03	123453	AC004971 Homo sapiens PAC cl
gb_pat:HS407A10	-	121.50	101.78	2.6e+03	33804	HS407A10 Human DNA sequence fr
gb_hcg:OSJN01012	-	121.50	89.94	1.2e+04	166307	OSJN01012 Oryza sativa chromo
gb_pat:AE006464	-	121.50	86.73	1.8e+04	256073	AE006464 Homo sapiens 16p13
gb_hcg:AC023841	-	121.00	89.69	1.2e+04	161297	AC023841 Homo sapiens chromo
gb_pat:AC005095	-	120.50	91.41	1.0e+04	119397	AC005095 Homo sapiens BAC cl
gb_hcg:AC095292	-	120.50	90.91	1.1e+04	128280	AC095292 Rattus norvegicus d

gb\_pat:U82668 - 119.50 92.99 8.1e+03 85132 1 U82668 Homo sapiens shox g  
gb\_hcg:AC008749.2 - 119.50 91.09 1.0e+04 110000 1 Continuation (3 of 5) of  
gb\_hcg:AC008749.3 - 119.50 91.09 1.0e+04 110000 1 Continuation (4 of 5) of  
gb\_hcg:AC026803 - 119.50 85.56 2.1e+04 231450 1 AC026803 Homo sapiens chr  
gb\_hcg:AP003936 - 117.50 87.45 1.7e+04 138459 1 AP003936 Oryza sativa chr  
seq\_name: gb\_pat:AX031287  
seq\_documentation\_block:  
LOCUS AX031287 596 bp DNA linear PAT 20-SEP-2000  
DEFINITION Sequence 9 from Patent W09914321.  
ACCESSION AX031287  
VERSION AX031287.1 GI:10278618  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 596)  
AUTHORS O'Reilly, L., Puthalakath, H., Adams, J., O'Connor, L., Cory, S.,  
Huang, D.C. and Strasser, A.  
TITLE Novel therapeutic molecules  
JOURNAL Patent: W0 9914321-A 9 25-MAR-1999;  
INST MEDICAL W & E HALL (AU); PUTHALAKATH HANSA (AU); REILLY  
LOHRAINE O (AU); ADAMS JERRY (AU); CONNOR LIAM O (AU); CORY  
SUZANNE (AU); HUANG DAVID C S (AU); STRASSER ANDREAS (AU)  
FEATURES  
source 1. 596  
/organism="unidentified"  
/db\_xref="taxon:32644"  
CDS  
1. >594  
/codon\_start=1  
/note="unnamed protein product"  
/protein\_id="CAC09657.1"  
/translation="MAKQSDVSDSECDREGLOPAERPPOLRPGAPTSLQTEPQGNP  
EGNHGGGSCPGSPQGLAPASPGFATRSLFIPMRSSLLSRSSGVSFSDTD  
RSPAPMSCKSTQTPSPQAPFNHYLSANASHMQAEAPADMRPEITWIAQELRRIGDEFN  
AYAVRRVFLNNTQAAEDHPRMVLRLRYIVLRVNRMH"  
BASE COUNT 145 a 175 c 146 g 130 t  
ORIGIN  
alignment\_scores:  
Quality: 1065.00 Length: 198  
Ratio: 5.379 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-508-832-10 x AX031287  
Align seg 1/1 to: AX031287 from: 1 to: 596  
1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17  
1 ATGGCAAGCAACCTTCGTATGATTAAGTCTAGTGTGACCGAAGGTAG 50  
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34  
51 ACAATGCGAGCTGGCGAGAGGCTCCCGACCTCAGACCTGGGGCCCTA 100  
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHsGlyGly 50  
101 CCTCCTACAGACAGAGCCACAGGTATCTCTGAAGCAATCAGGAGGT 150  
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProP 67  
151 GAAGGGGACAGTGGCCCCCAGCGAGCCCTCAGGGCCCGCTGGCCCCACC 200  
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeta 84  
201 TGGCAGCCCTGGCCCTTTTGGCTACCATCCCGCTTTCATCTTTATGA 250  
84 rgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheAsp 100

```

|||||
251 GAAGATCCTCCCTGCTGCTCGATCCTCAGTGGGTATTTCTCTTTTGAC 300
|||||
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
|||||
301 ACACAGAGGAGCCACGACCCATGAGTTGTGACAAATCAACACACCC 350
|||||
117 oSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134
|||||
351 AAGTCTCCTTTGGCAGGCTTCAACCACTATCTCAGTGCATGCTTCCA 400
|||||
134 eTarGlnAlaGluProAlaAspMetArgProGluIleTPrIleAlaGln 150
|||||
401 TGAGCAGGCTGAACCTGCAGATATGCGCCAGAGATATGATCGCCCA 450
|||||
151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaArgGva 167
|||||
451 GAGTTGGCGGCTATCGGAGACGAGTTTAACTGCTTACTATGCAAGGAG 500
|||||
167 IpeLeuAsnAsnTyrGlnAlaGluAspHisProArgMetValIleL 184
|||||
501 ATTTTGAATAATTACCAAGCAGCCAGACCCACCAAGATGCTTATCT 550
|||||
184 euArgLeuLeuArgTyrIleValArgLeuValTPrArgMetHis 198
|||||
551 TAGCACTGTACGTTACATGTGCGCTGGTGGAGATGCAT 594

```

seq\_name: gb\_un:AX031313

```

seq_documentation_block:
LOCUS AX031313 596 bp DNA linear UNA 20-SEP-2000
DEFINITION Sequence 9 from Patent WO9914321.
ACCESSION AX031313
VERSION AX031313.1 GI:10278641
KEYWORDS
SOURCE
ORGANISM

```

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

unidentified.

```

FEATURES
source
1..596
/organism="unidentified"
/db_xref="taxon:32644"
1..>594
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09662.1"
/db_xref="GI:10278642"
/translation="MAKQPSDVSSDCREGLOPAERPPOLRPGAPTSLOTPEPOGPNP
EGNHGSGDSCPHGSPGGLAPSPGPPATSPPLFIPMRSSLLSRSSSGYFSDTD
RSPAGKCDKSTQSPGCAFNHLYLSAMRQAEADMRPEIWIQAEKLRIGDEFN
AYVARRVELNNYQRAEDHPHVVILLRLVIVLWRKH"
BASE COUNT 145 a 175 c 146 g 130 t
ORIGIN

```

alignment\_scores:

```

Quality: 1065.00 Length: 198
Ratio: 5.379 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment\_block:

US-09-508-832-10 x AX031313

Align seg 1/1 to: AX031313 from: 1 to: 596

```

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyVar 17
|||||
1 ATGGCAAGCAACCTTCTGATGTAAGTTCTGAGTCCGACCGAGAGGTAG 50
|||||
17 GlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34
|||||
51 ACATTTGACGCTCGGAGAGGCTCCCGACGCTCAGACCTGGGGCCCTA 100
|||||
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
|||||
101 CCTCCTCTACAGACAGACCCACAAGGTAATCTCTGAAGGCAATCACGGAG 150
|||||
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
|||||
151 GAAGGGGACAGCTGCCCCCAGGAGCCCTCAGGGCCCGCTGGCCCCACC 200
|||||
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeta 84
|||||
201 TGCCAGCCCTGGCCCTTTTGTCTACAGATCCCGCTTTTTCATCTTTATGA 250
|||||
84 rArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 100
|||||
251 GAAGATCTCTCTCTCGATCCTCCAGTGGGTATTTCTCTTTTGAC 300
|||||
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
|||||
301 ACAGACAGGAGCCAGCCACCCATGAGTTGTGACAAATCAACACAAACCCC 350
|||||
117 oSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134
|||||
351 AAGTCTCCTTTGGCAGGCTTCAACCACTATCTCAGTGCATGCTTCCA 400
|||||
134 eTarGlnAlaGluProAlaAspMetArgProGluIleTPrIleAlaGln 150
|||||
401 TGAGCAGGCTGAACCTGCAGATATGCGCCAGAGATATGATCGCCCA 450
|||||
151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaArgGva 167
|||||
451 GAGTTGGCGGCTATCGGAGACGAGTTTAACTGCTTACTATGCAAGGAG 500
|||||
167 IpeLeuAsnAsnTyrGlnAlaGluAspHisProArgMetValIleL 184
|||||
501 ATTTTGAATAATTACCAAGCAGCCAGACCCACCAAGATGCTTATCT 550
|||||
184 euArgLeuLeuArgTyrIleValArgLeuValTPrArgMetHis 198
|||||
551 TAGCACTGTACGTTACATGTGCGCTGGTGGAGATGCAT 594

```

seq\_name: gb\_pr:AF032457

seq\_documentation\_block:

```

LOCUS AF032457 597 bp mRNA linear PRI 19-FEB-1998
DEFINITION Homo sapiens BimEL mRNA, complete cds.
ACCESSION AF032457
VERSION AF032457.1 GI:2895495
KEYWORDS
SOURCE
ORGANISM

```

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 597)

O'Connor, L., Strasser, A., O'Reilly, L.A., Hausmann, G., Adams, J.M.,

Cory, S. and Huang, D.C.

Bim: a novel member of the Bcl-2 family that promotes apoptosis

EMBO J. 17 (2), 384-395 (1998)

98094360

PUBMED

2 (bases 1 to 597)

O'Connor, L., Strasser, A., O'Reilly, L.A., Hausmann, G., Adams, J.M.,

Cory, S. and Huang, D.C.S.

Direct Submission

Submitted (03-NOV-1997)

Molecular Genetics of Cancer, The Walter &amp;

Eliza Hall Institute of Medical Research, PO Royal Melbourne

Hospital, Parkville, Victoria 3050, Australia

FEATURES

source  
1..597  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1..597

CDS  
/note="pro-apoptotic BH3-containing Bcl-2 family member"

/codon\_start=1  
/product="BimEL"  
/protein\_id="AAC39593.1"  
/db\_xref="GI:2895496"  
/translation="MAKOPSDVSECDREGLOPAERPPOLRCPADPTSLQTEPQGNP  
EGNHGEGSDSCPHSGPOGLAPASPCPEATSPLEPMRSSLRSSSGVSFSDTD  
RSPAPKSDKSTQTPSPCCOENHYLSAMASHROAPDMRPEITHIAELRIQDEEN  
ATYARVFLNNTQAEEDHPRWYLRLLRTIVRLVWRMH"  
BASE COUNT 146 a 175 c 146 g 130 t  
ORIGIN

alignment\_scores:  
Quality: 1065.00 Length: 198  
Ratio: 5.379 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-508-832-10 x AF032457 ..

Align seg 1/1 to: AF032457 from: 1 to: 597

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17  
|||||  
1 ATGGCAAGCAACCTTCTGATGTAAGTCTGAGTGTGACCGAGAGGTAG 50  
17 gGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34  
|||||  
51 ACRAATTCAGCCTCGGAGAGGCTCCCGACGCTCAGACCTGGGGCCCTA 100  
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50  
|||||  
101 CCTCCCTACAGACAGACCAAGGTAATCTCTGAAGGCAATCAGCGAGGT 150  
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67  
|||||  
151 GAAGGGGACAGTGCCTCCACGCGAGCCCTCAGGCGCCGCTGGCCGCCAC 200  
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPhePheMetA 84  
|||||  
201 TCCAGCCCTGGCCCTTTTGTACCGATCCCGCTTTTCATCTTTATGA 250  
84 rArgSerSerLeuSerArgSerSerSerGlyTyrPheSerPheAsp 100  
|||||  
251 GAAGATCCTCCCTGCTGCTCGATCCTCCAGTGGGTATTTCTTTTCAC 300  
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117  
|||||  
301 ACAGACAGAGCCGACGACCCCTGAGTGTGACAAATCAACACAAACCCC 350  
117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134  
|||||  
351 AAGTCTCTCTGCGGAGCCCTTCAACCACTATCTCAGTCGATGCGATGCTCA 400  
134 eArgGlnAlaGluProAlaAspMetArgProGluTlePrPleAlaGln 150  
|||||  
401 TCAGCGAGGCTGACCTGCGATATGCGCCAGAGATATGATGATGCGCCAA 450  
151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaArgArgVa 167  
|||||  
451 GAGTTGCCGGGTATCGGAGAGGAGTTTACGCTTACTATGCAAGGAGGGT 500  
167 lPheLeuAsnAsnTyrGlnAlaLaCluAspHisProArgMetValIleL 184  
|||||  
501 ATTTTGTGAATATTACCAAGCAGCGGAGACCAACCCACGAATGGTTATCT 550

184 euArgLeuLeuArgTyrIleValargLeuValTrpArgMetHis 198  
|||||  
551 TAGCACTGTTCAGTTACATTCCTGCCCTGGTGGAGAAATGCAT 594

seq\_name: gb\_pr:AB071197

seq\_documentation\_block:  
LOCUS AB071197 630 bp mRNA linear PRI 05-DEC-2001  
DEFINITION Homo sapiens mRNA for Bim-betal, complete cds.

ACCESSION AB071197

VERSION AB071197.1 GI:17351900

KEYWORDS Homo sapiens cDNA to mRNA.

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (sites)

AUTHORS U.M., Miyashita,T., Shikama,Y., Tadokoro,K. and Yamada,M.

TITLE Molecular cloning and characterization of six novel isoforms of

JOURNAL human Bim, a member of the proapoptotic Bcl-2 family(1)

PUBMED FEBS Lett. 509 (1), 135-141 (2001)

REFERENCE 2 (bases 1 to 630)

AUTHORS Miyashita,T.

TITLE Direct Submission

JOURNAL Submitted (06-SEP-2001) Toshiyuki Miyashita, National Children's

Medical Research Center, Department of Genetics, 3-35-31 Taishido,

Setagaya, Tokyo 154-8509, Japan (E-mail:tmlyashitaenchi.go.jp,

Tel:81-3-3414-8121(ex.2763), Fax:81-3-3414-3208)

FEATURES Location/Qualifiers

source 1..630

organism="Homo sapiens"

db\_xref="taxon:9606"

gene 1..408

/gene="bim-betal"

CDS 1..408

/gene="bim-betal"

/note="alternatively spliced: see also AccT

AB071195-AB071196, AB071198-AB071200"

/codon\_start=1

/product="Bim-betal"

/protein\_id="BAB78591.1"

/db\_xref="GI:17351901"

/translation="MAKOPSDVSECDREGLOPAERPPOLRCPADPTSLQTEPQGNP

EGNHGEGSDSCPHSGPOGLAPASPCPEATSPLEPMRSSLRSSSGVSFSDTD

RSPAPKSDKSTQTPSPCCOENHYLSAMASHROAPDMRPEITHIAELRIQDEEN"

BASE COUNT 157 a 186 c 152 g 135 t

ORIGIN

alignment\_scores:  
Quality: 1049.50 Length: 209  
Ratio: 5.301 Gaps: 1  
Percent Similarity: 94.737 Percent Identity: 94.737

alignment\_block:

US-09-508-832-10 x AB071197 ..

Align seg 1/1 to: AB071197 from: 1 to: 630

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17  
|||||  
1 ATGGCAAGCAACCTTCTGATGTAAGTCTGAGTGTGACCGAGAGGTAG 50  
17 gGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34  
|||||  
51 ACRAATTCAGCCTCGGAGAGGCTCCCGACGCTCAGACCTGGGGCCCTA 100  
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50  
|||||  
101 CCTCCCTACAGACAGACCAAGGTAATCTCTGAAGGCAATCAGCGAGGT 150  
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67

```

|||||
151 CAAGGGACAGCTGCCCCACGGACGCTCAGGGCCCGCTGGCCCCACC 200
|||||
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheAlaPheMeta 84
|||||
201 TGCCAGCCCTGGCCCTTTTGCTACCAAGATCCCGCTTTCATCTTTATGA 250
|||||
84 rArgSerSerLeuLeuSerArgSerSerSerGlyTyrrPheSerPheAsp 100
|||||
251 GAAGATCCTCCCTGCTGCTCGATCTCCAGTGGGTATTTCTCTTTTGAC 300
|||||
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
|||||
301 ACACAGAGGAGCCAGCACCCATGAGTTGTGACAAATCAACACAAACCCC 350
|||||
117 oSerProProCysGlnAlaPheAsnHisTyrrLeuSerAlaMet..... 131
|||||
351 AAGTCCCTCTGCCAGGCCCTTCAACCACTATCTCAGTGCATGGCTAACT 400
|||||
132 .....AlaSerMetArgGlnAlaGluPro 139
|||||
401 GGGACTAGAAACAGCTCCATCCACGCTTCCATGAGCGAGGCTGAACCT 450
|||||
140 AlaAspMetArgProGluIleThrPheAlaGlnGluLeuArgArgIleGl 156
|||||
451 CCACATATGCCCGCCACAGATATGATGCCCAAGAGTGGCGCGTATCGG 500
|||||
156 yAspGluPheAsnAlaTyrrTyrrAlaArgArgValPheLeuAsnAsnTyrg 173
|||||
501 AGACGAGTTTAAAGCTTACTATGCAAGGAGGGGTATTTTGAATAATATACC 550
|||||
173 lAlaAlaGluAspHisProArgMetValIleLeuArgGluLeuArgTyrr 189
|||||
551 AAGCAGCGGAAGACCCACCCACCAATGTTATCTTACGACTGTTAGGTAC 600
|||||
190 lIleValArgLeuValTrpArgMetHis 198
|||||
601 ATTGTCCGCTGGTGTGGAGATGAT 627

```

seq\_name: 9b\_pr:AB071195

seq\_documentation\_block:

LOCUS AB071195 722 bp mRNA linear PRI 05-DEC-2001

DEFINITION Homo sapiens mRNA for Bim-alpha1, complete cds.

ACCESSION AB071195

VERSION AB071195.1 GI:17351896

KEYWORDS

SOURCE Homo sapiens cDNA to mRNA.

ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE Molecular cloning and characterization of six novel isoforms of

human Bim, a member of the proapoptotic Bcl-2 family(1)

JOURNAL FEBS Lett. 509 (1), 135-141 (2001)

PUBMED 11734221

REFERENCE 2 (bases 1 to 722)

AUTHORS

TITLE Direct Submission

Submitted (06-SEP-2001) Toshiyuki Miyashita, National Children's

Medical Research Center, Department of Genetics; 3-35-31 Taishido,

Setagaya, Tokyo 154-8509, Japan (E-mail: tmiyashita@nch.go.jp,

Tel:81-3-3414-8121(ex.2763), Fax:81-3-3414-3208)

FEATURES

source 1..722

/organism="Homo sapiens"

/db\_xref="taxon:9606"

gene 1..510

/gene="bim-alpha1"

cds 1..510

/gene="bim-alpha1"

/note="alternatively spliced: see also Acc#

```

AB071196-AB071200*
/codon_start=1
/product="bim-alpha1"
/protein_id="BAB8589.1"
/db_xref="GI:17351897"
/translation="MAKOPSDVSECDREGLOAERPPOLRPGAPTSLOTPEOGNP
EGNHGGEGDSCPHGSPGAPATPSPLFIMRRSSLLSRSSSGYFFSFDTD
RSPAPMSCDKSTQTPSPQCAFNHILSAMASHMQAEADMRPEIWIAGELRRIGDEFN
AYYARRLEK"
BASE COUNT 180 a 190 c 193 g 159 t
ORIGIN

```

alignment\_scores:

Quality: 1010.50 Length: 240

Ratio: 5.104 Gaps: 2

Percent Similarity: 82.500 Percent Identity: 82.083

alignment\_block:

US-09-508-832-10 x AB071195

Align seg 1/1 to: AB071195 from: 1 to: 722

```

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
|||||
1 ATGGCCAAAGCAACCTTCTGATGAAGTTCTGAGTGTGACCGAGAGGTAG 50
|||||
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34
|||||
51 ACAATTCGAGCCTCGCGAGAGGCTCCCGAGCTCAGACCTGGGGCCCTA 100
|||||
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
|||||
101 CCTCCTACACAGAGCCACAGGTATCTCTCAAGGCAATCACGGAGGT 150
|||||
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
|||||
151 GAAGGGAGACCTGCCCCACGGAGCCTCAGGGCCCGCTGGCCCCACC 200
|||||
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheAlaPheMeta 84
|||||
201 TGCCAGCCCTGGCCCTTTTGCTACCAAGATCCCGCTTTTCATCTTTATGA 250
|||||
84 rArgSerSerLeuLeuSerArgSerSerSerGlyTyrrPheSerPheAsp 100
|||||
251 GAAGATCCTCCCTGCTGCTCGATCTCCAGTGGGTATTTCTCTTTTGAC 300
|||||
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
|||||
301 ACACAGAGGAGCCAGCACCCATGAGTTGTGACAAATCAACACAAACCCC 350
|||||
117 oSerProProCysGlnAlaPheAsnHisTyrrLeuSerAlaMetAlaSerM 134
|||||
351 AAGTCCCTCTTGCCAGGCCCTTCAACCACTATCTCAGTGCATGGCTTCCA 400
|||||
134 eArgGlnAlaGluProAlaAspMetArgProGluIleThrPheAlaGln 150
|||||
401 TGAGCGAGGCTGAACCTGCGAGATATGCGCCACAGAGATATGGATGCCCA 450
|||||
151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrrTyrrAla..... 164
|||||
451 GAGTTGCGGCGGTATCGGAGACGAGTTTAAACGCTTACTATGCAAGGAGGT 500
|||||
164 .....Arg...Arg...ArgValPheLeuAsnAsnTyrrGlnAlaAla 175
|||||
501 ACAGAAATAGAGGAAGTTGTCGTGTAGTTCTCATGCTATTTCAGTCCACCTA 550
|||||
164 ..... 164
551 AGGGCAGTGGGGAAGCGCTTTTCAGACCGGAGCCTGCGAGGCTGAATCCTTGA 600
|||||
165 .....Arg...Arg...ArgValPheLeuAsnAsnTyrrGlnAlaAla 175
|||||

```



601 AGGAGGAGGTGAGAGAGCAGCAGGTATTTTGAATAATTACCAAGCAGCC 650  
 176 GluAspHisProArgMetValIleLeuArgLeuLeuArgTyrIleValar 192  
 651 GAAGACCAACCCACCAAGTGTATCTACGACTGTTACGTTACATTTCTCCG 700  
 192 GluValTrpArgMetHis 198  
 701 CTTGGTGTGGAGATGCAT 719

seq\_name: gb\_pat:AX031283

seq\_documentation\_block:  
 LOCUS AX031283 590 bp DNA linear PAT 20-SEP-2000  
 DEFINITION Sequence 5 from Patent W09914321.  
 ACCESSION AX031283  
 VERSION AX031283.1 GI:10278614

KEYWORDS  
 SOURCE

ORGANISM  
 UNIDENTIFIED

REFERENCE  
 1 (bases 1 to 590)

AUTHORS  
 O'Reilly, L., Puthalakath, H., Adams, J., O'Connor, L., Cory, S.,  
 Huang, D.C. and Strasser, A.

TITLE  
 Novel therapeutic molecules

JOURNAL  
 Patent: WO 9914321-A 5 25-MAR-1999;

LORRAINE O (AU); ADAMS JERRY (AU); PUTHALAKATH HANSA (AU); REILLY  
 SUZANNE (AU); HUANG DAVID C S (AU); STRASSER ANDREAS (AU)

FEATURES

source  
 1..590  
 /organism="unidentified"  
 /db\_xref="taxon:32644"

CDS  
 1..>587  
 /note="unnamed protein product"

/codon\_start=1  
 /protein\_id="CAC09655.1"

/db\_xref="GI:10278615"

/translation="MAKQPSDVSSCCDREGGQLOAERPQLRPGAPTSLQTEPOGPN  
 DGGDRCPHSGPOGLAPASPATPSPLFIFVRRSSLSRSSGYSFSDTRSPA  
 PMSCDKSTQTPSPPCQAFNHYSAMASIROQDEPDLRPEIRIAQLRRIGDEFNET  
 YTRRVFANDYREADHPQVILQLRIFRLVWRR"

BASE COUNT 137 a 178 c 150 g 125 t

ORIGIN

alignment\_scores:  
 Quality: 908.00 Length: 200  
 Ratio: 4.935 Gaps: 2

Percent Similarity: 92.000 Percent Identity: 86.500

alignment\_block:

US-09-508-832-10 x AX031283 ..

Align seg 1/1 to: AX031283 from: 1 to: 590

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17

1 ATGGCCACAGCACCTCTGATGTAAAGTTCTGAGTGTGACAGAGAGAGTGG 50

17 gGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34

51 ACATTTGCAGCTGCTGAGAGGCTCCCTCCAGCTCAGGCTGGGGCCCTA 100

34 hrSerLeuGlnThrGluProGlnCysProGluGlyAsnHisGlyGly 50

101 CTTCCCTACACAGACACCGACGAGTANTCCGAC.....GGC 138

51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67

139 GAAGGGGACCGCTGCCCCACGCGCCCTCAGGGCCCGCTGGCCCCACC 188

67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeLa 84

189 GCCACGCCCTGGCCCTTTTCTACAGATCCCACTTTTCATCTTTGTGA 238

84 rgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheAsp 100

239 GAGATCTTCTGCTGCTCCCGTCTCTCCAGTGGGTATTTCTCTTTTGAC 288

101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117

289 ACAGACAGGAGCCCGCACCATGAGTTGTGACAAAGTCAACACAAACCCC 338

117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134

339 AAGTCTCTTTCAGCGCTTCAACCACTATCTCAGTGCATGGCTTCCA 388

134 eArgGlnAla.....GluProAlaAspMetArgProGluIleTrpIle 148

389 TAGCACAGCTCTCAGGAGGAACTGAGATCTGCCCGCGAGATACGGATT 438

149 AlaGlnGlnLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaAr 165

439 GCACAGGAGCTCCCGCGGATCGGACAGCTTCAACGAAACTTACACAAG 488

165 gArgValPheLeuAsnAsnTyrGlnAlaAlaGluAspHisProArgMetV 182

489 GAGGTGTTTTCGAAATGATTACCGGAGCTGAAGACCAACCCCTCAATGG 538

182 aIleLeuArgLeuLeuArgTyrIleValArgLeuValTyrArgMetHis 198

539 TTATCTTACAACTGTTACGCTTTATCTTCCTGCTGGTATGAGAGGCAT 588

seq\_name: gb\_un:AX031309

seq\_documentation\_block:

LOCUS AX031309 590 bp DNA linear UNA 20-SEP-2000

DEFINITION Sequence 5 from Patent W09914321.

ACCESSION AX031309

VERSION AX031309.1 GI:10278637

KEYWORDS  
 SOURCE

ORGANISM  
 UNIDENTIFIED

REFERENCE  
 1 (bases 1 to 590)

AUTHORS  
 O'Reilly, L., Puthalakath, H., Adams, J., O'Connor, L., Cory, S.,  
 Huang, D.C. and Strasser, A.

TITLE  
 Novel therapeutic molecules

JOURNAL  
 Patent: WO 9914321-A 25-MAR-1999;  
 INST MEDICAL W & E HALL (AU); PUTHALAKATH HANSA (AU); REILLY  
 LORRAINE O (AU); ADAMS JERRY (AU); CONNOR LIAM O (AU); CORY  
 SUZANNE (AU); HUANG DAVID C S (AU); STRASSER ANDREAS (AU)

FEATURES  
 Location/Qualifiers

source  
 1..590  
 /organism="unidentified"

/db\_xref="taxon:32644"

1..>587  
 /note="unnamed protein product"

/codon\_start=1  
 /protein\_id="CAC09660.1"

/db\_xref="GI:10278638"

/translation="MAKQPSDVSSCCDREGGQLOAERPQLRPGAPTSLQTEPOGPN  
 DGGDRCPHSGPOGLAPASPATPSPLFIFVRRSSLSRSSGYSFSDTRSPA  
 PMSCDKSTQTPSPPCQAFNHYSAMASIROQDEPDLRPEIRIAQLRRIGDEFNET  
 YTRRVFANDYREADHPQVILQLRIFRLVWRR"

BASE COUNT 137 a 178 c 150 g 125 t

ORIGIN

alignment\_scores:

Quality: 908.00 Length: 200

Ratio: 4.935 Gaps: 2

Percent Similarity: 92.000 Percent Identity: 86.500

alignment\_block:

US-09-508-832-10 x AX031309

Align seg 1/1 to: AX031309 from: 1 to: 590

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17  
 1 ATGGCCAGGACCTTCTGATGTAAGTCTCAGTGTGACAGAGAGGTGG 50  
 17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPro 34  
 51 ACAATTGACGCTGCTGAGAGGCTCCCGAGCTCAGGCTGGGCCCTA 100  
 34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGly 50  
 101 CCTCCCTACAGACAGACCCGAGGTAAATCCGAC.....GGC 138  
 51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPro 67  
 139 GAAGGGAGACCGCTGCCCGCCAGCGACGCTCAGGCGCCGCTGGCCCA 188  
 67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMet 84  
 189 GCCAGGCTGCGCTTTGCTACAGATCCCGACCTTTTCATCTTTGTA 238  
 84 rgArgSerLeuLeuSerArgSerSerGlyTyrPheSerPheAsp 100  
 239 GAAGATCTTCTGCTGCTCCGCTCCTCAGTGGGTATTTCTCTTTGAC 288  
 101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThr 117  
 289 ACAGACAGGAGCGCGCACCATGAGTTGTGACAAGTCAACACAAACCC 338  
 117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSer 134  
 339 AAGTCTCTTGGCAGGCTTCAACCACTATCTCAGTGCATGGCTTCCA 388  
 134 etArgGlnAla.....GluProAlaAspMetArgProGluIleTrp 148  
 389 TACGACAGCTCAGGAGACCTCGAGATCTGCCCGCGGATACGGATT 438  
 149 AlaGlnGluLeuArgArgIleGlyAspGluPheAsnAlaTyrAlaAr 165  
 439 GCACAGGAGCTGGCGGGATCGGAGACGAGTTCACAGAACTTACACA 488  
 165 gArgValPheLeuAsnAspTyrGlnAlaAlaGluAspHisProArgMet 182  
 489 GAGGCTCTTGCATGATTACCGGAGCTGAGAGACCCCTCAATGG 538  
 182 alIleLeuArgLeuLeuArgTyrIleValArgLeuValTyrArgMetHis 198  
 539 TTATCTTACAACTGTTACGCTTATCTCTCCGCTGCTGATGAGAAGCAT 588

seq\_name: gb\_ro:AF032459

seq\_documentation\_block:

LOCUS AF032459 591 bp mRNA linear ROD 19-FEB-1998

DEFINITION Mus musculus BimEL mRNA, complete cds.

ACCESSION AF032459

VERSION AF032459.1 GI:2895499

KEYWORDS house mouse.

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 591)

O'Connor, L., Strasser, A., O'Reilly, L.A., Hausmann, G., Adams, J.M.,

Cory, S. and Huang, D.C.

Bim, a novel member of the Bcl-2 family that promotes apoptosis

EMBO J 17 (2), 384-395 (1998)

MEDLINE 98094360

PUBMED 9430630

REFERENCE 2 (bases 1 to 591)

O'Connor, L., Strasser, A., O'Reilly, L.A., Hausmann, G., Adams, J.M.,

Cory, S. and Huang, D.C.S.

Direct Submission

TITLE

Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter &amp; Eliza Hall Institute of Medical Research, PO Royal Melbourne Hospital, Parkville, Victoria 3050, Australia

FEATURES

Location/Qualifiers

source

1..591

/organism="Mus musculus"

/db\_xref="taxon:10090"

1..591

/note="pro-apoptotic Bcl-2 family member"

/codon\_start=1

/product="BimEL"

/protein\_id="AAC40029.1"

/db\_xref="GI:2895500"

/translation="MAKOPSDVSSCDREGGQLQAPRRPQLRPGAPTSLSLQTEPGNP  
 DGEGRCHGSPQGLPAPSPFPATRSPLFIVRRSLLSRSSSGYFSDTDRSPA  
 PMSCKSTQTFSPPCQAFNFHLSAMASIRQSEBEDLRPEIRIAQLRRIQIDBFNET  
 YTRRVFANDYREADHPOMVILQLRIFRLVYRRH"

BASE COUNT 138 a 178 c 150 g 125 t

ORIGIN

alignment\_scores:

Quality: 908.00 Length: 200

Ratio: 4.935 Gaps: 2

Percent Similarity: 92.000 Percent Identity: 86.500

alignment\_block:

US-09-508-832-10 x AF032459

Align seg 1/1 to: AF032459 from: 1 to: 591

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17

1 ATGGCCAGGACCTTCTGATGTAAGTCTCAGTGTGACAGAGAGGTGG 50

17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPro 34

51 ACAATTGACGCTGCTGAGAGGCTCCCGAGCTCAGGCTGGGCCCTA 100

34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGly 50

101 CCTCCCTACAGACAGACCCGAGGTAAATCCGAC.....GGC 138

51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPro 67

139 GAAGGGAGACCGCTGCCCGCCAGCGACGCTCAGGCGCCGCTGGCCCA 188

67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMet 84

189 GCCAGGCTGCGCTTTGCTACAGATCCCGACCTTTTCATCTTTGTA 238

84 rgArgSerLeuLeuSerArgSerSerGlyTyrPheSerPheAsp 100

239 GAAGATCTTCTGCTGCTCCGCTCCTCAGTGGGTATTTCTCTTTGAC 288

101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThr 117

289 ACAGACAGGAGCGCGCACCATGAGTTGTGACAAGTCAACACAAACCC 338

117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSer 134

339 AAGTCTCTTGGCAGGCTTCAACCACTATCTCAGTGCATGGCTTCCA 388

134 etArgGlnAla.....GluProAlaAspMetArgProGluIleTrp 148

389 TACGACAGCTCAGGAGACCTCGAGATCTGCCCGCGGATACGGATT 438

149 AlaGlnGluLeuArgArgIleGlyAspGluPheAsnAlaTyrAlaAr 165

439 GCACAGGAGCTGGCGGGATCGGAGACGAGTTCACAGAACTTACACA 488



dimerization with diverse antiapoptotic Bcl-2 members  
 Mol. Endocrinol. 12 (9), 1432-1440 (1998)  
 98400436  
 2 (bases 1 to 591)  
 Hsu, S.Y. and Hsueh, A.J.W.  
 Direct Submission  
 Submitted (15-MAY-1998) GYN/OB, Stanford University, MSOB S385,  
 Stanford, CA 94305, USA  
 Location/Qualifiers  
 1..591  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 1..591  
 /codon\_start=1  
 /product="Bcl-2 related ovarian death gene product BOD-L"  
 /protein\_id="AAC23595.1"  
 /db\_xref="GI:3228570"  
 /translation="MAKQSDVNSCEDREGGLOPAERPPOLRPGAPTSLOTESQGNP  
 DGGDRCPHSGPGLAPSPATRSPLFIYVRRSSLSRSSSGYFSFDTRSPA  
 PHSCDKSTOTPPPCQAFNHYSAMASTROSEPEEDLPEIRIAQLRRIGDEFNET  
 YTRAFANDYREADHPQMVILQLLRFILVWRRH"  
 139 a 177 c 153 g 122 t  
 BASE COUNT  
 ORIGIN

alignment\_scores:  
 Quality: 893.00 Length: 200  
 Ratio: 4.880 Gaps: 2  
 Percent Similarity: 91.500 Percent Identity: 85.000

alignment\_block:  
 US-09-508-832-10 x AF065433 ..  
 Align seg 1/1 to: AF065433 from: 1 to: 591

1 MetAlaLysGlnProSerAspValSerGluCysAspArgGluGlyAr 17  
 1 ATGCCAAGCAACCTCTGATGTAATCTGAGTGTACAGAGAGGTGG 50  
 17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34  
 51 ACAATTGACCTGCTGACAGAGCCCTCCCACTGAGTGTGGGCCCTTA 100  
 34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50  
 101 CCTCCCTACAGACAGAAATGCCAAGGTAAATCCCGAC.....GGC 138  
 51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67  
 139 GAAGGGACCCGCTGCCCAACGACAGCCCTCAGGGCCGCTGGCCCAAC 188  
 67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeta 84  
 189 GCCCAGCCCTGGTCTTTGGTACCAGATCCCACTTTTCATCTTTGTGA 238  
 84 rGatSerSerLeuLeuSerArgSerSerGlyTyPheSerPheAsp 100  
 239 GAAGATCTCTGCTGTGCTCCGCTCCTCCAGTGGGTATTTCTCTTTGAC 288  
 101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117  
 289 ACAGACAGAGAGCCCGCACCATGATGCTGACAACTCACACACACACCC 338  
 117 oSerProProCysGlnAlaPheAsnHisTyLeuSerAlaMetaSerM 134  
 339 AAGTCCCTCTGGCAGCCCTTCAACCATTTATCTCAGTGCATGCTGCCA 388  
 134 etArgGlnAla.....GluProAlaAspMetArgProGluIleT-pile 148  
 389 TAAGCCAGTCTCAGGAGGAACCTGAAGATCTGGCCCCAGAGATACGATC 438  
 149 AlaGlnLeuArgArgIleGlyAspGluPheAsnAlaTyTyAlaAr 165  
 1 ATGCGAAGCAACCTCTGATGTAAGTCTGAGTGTGACCCAGAGAGGTAG 50  
 17 gClnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34  
 51 ACATATGACGCTCGGAGAGAGCTTCCCGAGCTCAGACCTGGGGCCCTTA 100

439 GCACAGGAGCTGCGCGGAGCGGAGAGAGTTCATGAGACTTACACGAG 488  
 165 qArgValPheLeuAsnTyrGlnAlaAlaGluAspHisProArgMetV 182  
 489 GAGGCGCTTTCGAACAGATTACCGAGGCGGAGACCCACCCCAATGG 538  
 182 alileLeuArgLeuLeuArgTyrIleValArgLeuValTrpArgMethis 198  
 539 TTATCTTAAACTGTTAGATTATCTTCGTCGTGCTGTGAGAGAGGCAC 588  
 seq\_name: gb\_pr:AB071198  
 seq\_documentation\_block:  
 LOCUS AB071198 493 bp mRNA linear PRI 05-DEC-2001  
 DEFINITION Homo sapiens mRNA for Bim-beta2, complete cds.  
 ACCESSION AB071198  
 VERSION AB071198.1 GI:17351902  
 KEYWORDS  
 SOURCE Homo sapiens cDNA to mRNA.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (sites)  
 AUTHORS U.M., Miyashita, T., Shikama, Y., Tadokoro, K. and Yamada, M.  
 TITLE Molecular cloning and characterization of six novel isoforms of  
 human Bim, a member of the proapoptotic Bcl-2 family(1)  
 JOURNAL FEBS Lett. 509 (1), 135-141 (2001)  
 PUBMED 11734221  
 REFERENCE 2 (bases 1 to 493)  
 AUTHORS Miyashita, T.  
 TITLE Direct Submission  
 JOURNAL Submitted (06-SEP-2001) Toshiyuki Miyashita, National Children's  
 Medical Research Center, Department of Genetics, 3-35-31 Taisshido,  
 Setagaya, Tokyo 154-8509, Japan (E-mail: tmiyashita@nch.go.jp,  
 Tel:81-3-3414-8121(ex-2763), Fax:81-3-3414-3208)

FEATURES  
 source  
 1..493  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 1..408  
 /gene="bim-beta2"  
 1..408  
 /gene="bim-beta2"  
 /product="Bim-beta2"  
 /protein\_id="BAB78502.1"  
 /db\_xref="GI:17351903"  
 /translation="MAKQSDVNSCEDREGGLOPAERPPOLRPGAPTSLOTESQGNP  
 EGNHGGEDSCPHSGPGLAPSPATRSPLFIYVRRSSLSRSSSGYFSFDTRD  
 RSPAPMSCDKSTOTPPPCQAFNHYSAMGTF"  
 BASE COUNT 119 a 151 c 114 g 109 t  
 ORIGIN

alignment\_scores:  
 Quality: 837.50 Length: 199  
 Ratio: 5.138 Gaps: 2  
 Percent Similarity: 81.910 Percent Identity: 81.407

alignment\_block:  
 US-09-508-832-10 x AB071198 ..  
 Align seg 1/1 to: AB071198 from: 1 to: 493

1 MetAlaLysGlnProSerAspValSerGluCysAspArgGluGlyAr 17  
 1 ATGCCAAGCAACCTCTGATGTAAGTCTGAGTGTGACCCAGAGAGGTAG 50  
 17 gClnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34  
 51 ACATATGACGCTCGGAGAGAGCTTCCCGAGCTCAGACCTGGGGCCCTTA 100



```

source          1. 416
                /organism="unidentified"
                /db_xref="taxon:32644"
                1. >414
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAC09656.1"
                /db_xref="GI:10278617"
                /translation="MAKQPSDVSSDCREGROLOPAERPPQLRPGAPTSIQTPEODRS
                PAPMSCDKSTQTPSPPCQAFNHYSAMASMRQAEADMRPEIWIQAELRRIGDEFNAY
                YARRVFLNYYQAAEDHPRMVILRLRYIVRLVRMH"
BASE COUNT      113 a 113 c 103 g 87 t
ORIGIN
alignment_scores:
  Quality: 696.00      Length: 198
  Ratio: 5.043        Gaps: 1
  Percent Similarity: 69.697   Percent Identity: 69.697
alignment_block:
US-09-508-832-10 x AX031285 ..
Align seg 1/1 to: AX031285 from: 1 to: 416
1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
|||||
1 ATGGCAAGGCAACCTTCGTATGTAAGTCTGTAGTGTGACCGAGAAGGTAG 50
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34
|||||
51 ACAATTGCAGCTGCGGAGAGGCTCCCGACGTCAGACCTGGGGCCCTA 100
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
|||||
101 CCTCCTACAGACAGAGCCACAA..... 123
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
123 ..... 123
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeTA 84
123 ..... 123
84 rArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheAsp 100
123 ..... 123
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
124 ...GACAGGAGCCCGACGCCCATGAGTTGTGTGACAAATCAACACAAACCCC 170
117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134
|||||
171 AAGTCCTCCTTGCAGGCTTCAACACACTATCTCAGTGCATGCGCTTCCA 220
134 eArgGlnAlaGluProAlaAspMetArgProGluIleTyrIleAlaGln 150
|||||
221 TGAGCAGGCTGACCTTCAGATATGCGCCAGAGATATGATGCGCCCAA 270
151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaArgArgVa 167
|||||
271 CAGTTGGCGGTATCGGAGACGAGTTTACGCTTACTATGCAAGGAGGTT 320
167 lPheLeuAsnAsnTyrGlnAlaAlaGluAspHisProArgMetValIleL 184
|||||
321 ATTTTTCANATNTACCAAGCAGCCGACCCACCCACCAATGTTATCT 370
184 euArgLeuLeuArgTyrIleValArgLeuValTrpArgMetHis 198
|||||
371 TACGACTGTTACGTTACATGTCGCCCTGCTGGTGTGGAGATGCAT 414

```

```

seq_name: gb_un:AX031311
seq_documentation_block:
LOCUS AX031311 416 bp DNA linear UNA 20-SEP-2000
DEFINITION Sequence 7 from Patent WO9914321.
ACCESSION AX031311
VERSION AX031311.1 GI:10278639
KEYWORDS
SOURCE
ORGANISM
unidentified.
unclassified.
REFERENCE
1 (bases 1 to 416)
O'Reilly,L., Ruthalakath,H., Adams,J., O'Connor,L., Cory,S.,
Huang,D.C. and Strasser,A.
Novel therapeutic molecules
Patent: WO 9914321-A 25-MAR-1999;
INST MEDICAL W & E HALL (AU); PUTHALAKATH HANSA (AU); REILLY
LORRAINE O (AU); ADAMS JERRY (AU); CONNOR LIAM O (AU); CORY
SUZANNE (AU); HUANG DAVID C S (AU); STRASSER ANDREAS (AU)
FEATURES
Location/Qualifiers
1. 416
/organism="unidentified"
/db_xref="taxon:32644"
CDS
1..>414
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09661.1"
/db_xref="GI:10278640"
/translation="MAKQPSDVSSDCREGROLOPAERPPQLRPGAPTSIQTPEODRS
PAPMSCDKSTQTPSPPCQAFNHYSAMASMRQAEADMRPEIWIQAELRRIGDEFNAY
YARRVFLNYYQAAEDHPRMVILRLRYIVRLVRMH"
BASE COUNT 113 a 113 c 103 g 87 t
ORIGIN
alignment_scores:
  Quality: 696.00      Length: 198
  Ratio: 5.043        Gaps: 1
  Percent Similarity: 69.697   Percent Identity: 69.697
alignment_block:
US-09-508-832-10 x AX031311 ..
Align seg 1/1 to: AX031311 from: 1 to: 416
1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
|||||
1 ATGGCAAGGCAACCTTCGTATGTAAGTCTGTAGTGTGACCGAGAAGGTAG 50
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34
|||||
51 ACAATTGCAGCTGCGGAGAGGCTCCCGACGTCAGACCTGGGGCCCTA 100
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
|||||
101 CCTCCTACAGACAGAGCCACAA..... 123
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
123 ..... 123
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeTA 84
123 ..... 123
84 rArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheAsp 100
123 ..... 123
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
124 ...GACAGGAGCCCGACGCCCATGAGTTGTGTGACAAATCAACACAAACCCC 170

```

1	MetAlaIylsGlnProSerAspValSerSerGluCysAspArgGluGlyArg	17
1	ATGGCAAGACCACTTCGATGTAAGTTCTGAGTGTGACCCAGGAAGGTAG	50
17	GcInLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT	34
51	ACAATTCGACGCTGGGGAGAGCCCTCCCCAGCTCAGACCTGGGGCCCCTA	100
34	hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly	50
101	CCTCCCTACAGACAGAGCCACAA	123
51	GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr	67
123	.....	123
67	OAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMetA	84
123	.....	123
84	rgArgSerSerLeuLeuSerArgSerSerSerGlyTyPheSerPheAsp	100
123	.....	123
101	ThrAspArgSerProAlaPrometSerCysAspLysSerThrGlnThrPr	117
124	...GACAGGAGCCAGCACCCATGAGTGTGTGACAATCAACACAAACCCC	170
117	oSerProProCysGlnAlaPheAsnHisTyLeuSerAlaMetAlaSerM	134
171	AAGTCTCCTTGCCAGGCGTTCAACCACTATCTCTAGTGCATGCTTCCA	220
134	eArgGlnAlaGluProAlaAspMetArgProGluTleTTrPleAlaGln	150
221	TGAGGAGGCTGAACCTGCAGATATGCGCCCCAGAGATATGGATCGCCCAA	270
151	GluLeuArgArgTleGlyAspGluPheAsnAlaTyTyAlaArgArgVa	167
271	GAGTTGCGGCGTATCGGAGCAGGTTTAAACCCCTTACTATGACAGAGGGT	320
167	lPheLeuAsnAsnTyGlnAlaAlaGluAspHisProArgMetValIleL	184
321	ATTTTGTAAATTAATACAGACGCGGAAGACACCCAGCAATGTTATCT	370
184	eArgLeuLeuArgTyTyrIleValArgLeuValTrpArgMetHis	198
371	TACCACTGTTACGTTACATGTCGCGCTGCTGTGGAGAAATGCAT	414

seq documentation block:

ACCESSION AF032458

VERSION  
KEYWORDS  
AF032438.1 GI:2093497

SOURCE	ORGANISM	human.	human

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Artiodactyla; Bovidae; Caprinae; Caprini; *Ovis montanus*

REFERENCE 1 (bases 1 to 417)

**Cory, S. and Huang, D.C.**

JOURNAL EMBO J. 17 (2), 384-399

PUBMED 94 30630

**AUTHORS**  
O'Connor, L.; Strasser,

TITLE  
Direct submission  
Cory, S. and Huang, D.C.

JOURNAL  
SUBMITTED  
11-23-2011  
INST 004070

Hospital, parkville, v

source 1. .417

/db\_xref="tax

```

/note="pro-ap

```

```

/product="Bim
/coach_star

```

/protein\_id=

/translation-  
nadmcsdpcstcd

YARRVELNNYQAA

**ORIGIN**

**alignment\_scores:**

Ratio: 5.043

REFERENCE: 03-037 PER

alignment\_block: 00000000-0000-0000-0000-000000000000  
ms-09-508-832-10 x 15033459

4  
C  
6  
7  
C  
C  
E  
4  
  
4  
5  
7  
1  
6  
5  
2

OM of: US-09-508-832-10 to: N\_Geneseq\_032802.\* out\_format : pfs  
 Date: Aug 7, 2002 10:40 PM  
 About: Results were produced by the GenCore software, version 4.5.  
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL-frame: p2n.model -DEV-xlh  
 -Q/cgn2.1/USPTO.spool/US09508832/runat\_06082002\_103612\_2816/app\_query.fasta\_1.259  
 -DB-N\_Geneseq\_032802 -OFMT=fastcap -SUFFIX=ring -GAPOP=12.000  
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
 -GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
 -GAPOP=6.000 -GAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomung2  
 -TRANS=human0.cdi -LIST=45 -DOCAUIGN=200 -THR\_SCORE=pct  
 -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
 -USER=US09508832@cgn1\_1\_0 -NCPU=6 -ICPU=3 -LONGLOG  
 -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-508-832-10  
 Query length: 198  
 Database: N\_Geneseq\_032802.\*  
 Database sequences: 1736436  
 Search time (sec): 211.310000

score\_list:

Sequence	Strd Orig	ZScore	EScore	Len	Document
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA24997		1065.00	1251.99	2.2e-61	59
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA24997		908.00	1068.41	3.7e-51	58
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA24996		696.00	823.23	1.7e-37	41
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA24994		589.00	697.94	1.6e-30	42
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA565350		497.00	589.94	1.6e-24	33
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA24993		406.00	485.79	1.0e-18	34
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA189284		123.00	153.83	3.21	3
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AA35620		119.50	113.58	559.30	32
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA207131		113.00	125.78	117.05	2
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA207132		113.00	125.77	117.26	2
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA207133		113.00	125.76	117.27	2
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA24943		112.00	135.09	35.44	7
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA584424		112.00	125.96	114.40	2
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA58472		112.00	109.27	973.18	19
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA166848		111.50	123.23	162.38	3
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA166832		111.50	122.51	177.94	3
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA166831		111.50	122.46	179.26	3
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA166833		111.50	122.46	179.26	3
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA166853		111.50	122.46	179.26	3
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA32265		111.50	121.76	196.03	37
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA488200		111.50	121.76	196.03	37
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA571257		111.00	121.76	196.03	37
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA14569		111.00	125.24	110.36	1
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA76509		111.00	125.43	122.40	21
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA320163		109.00	129.53	72.31	98
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA320163		109.00	100.88	2.9e+03	34
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA50085		108.50	127.04	99.51	12
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA58471		108.50	95.88	5.4e+03	58
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA65017		107.50	128.86	78.88	8
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA60529		107.50	124.02	146.58	8
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA505497		107.50	121.57	200.86	2
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA17441		107.50	120.56	228.56	2
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA16913		107.50	116.33	393.32	3
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA64966		107.50	180.29	4.0e+04	3
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AA34609		107.00	130.61	62.98	64
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA113344		107.00	130.61	62.98	64
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AA79322		107.00	126.76	103.16	10
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AA34608		107.00	123.44	157.90	15
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA211371		107.00	123.16	163.86	16

/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA126749 - 107.00 115.02 455.15  
 /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AA52403 - 107.00 113.52 563.98  
 /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ5933 - 106.50 125.16 126.75  
 /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAA6482 - 106.50 124.50 137.91  
 /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AA578921 - 106.50 122.71 173.58

seq\_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AA24997

seq\_documentation\_block:

ID AAX24997 standard; cDNA: 596 BP.  
 XX  
 AC AAX24997;  
 XX  
 DT 05-JUL-1999 (first entry)  
 XX  
 DE Human Bcl-2 interacting mediator of cell death Bim-EL cDNA.  
 XX  
 KW Bim-EL; Bcl-2 interacting mediator of cell death; apoptosis;  
 KW cell cycle; human; cancer; autoimmune disease;  
 KW degenerative disease; therapy; contraceptive; splice variant;  
 KW isoform; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09914321-A1.  
 XX  
 PD 25-MAR-1999.  
 XX  
 PF 17-SEP-1998; 98WO-AU00772.  
 XX  
 PR 24-SEP-1997; 97AU-0009373.  
 PR 17-SEP-1997; 97AU-0009263.  
 XX  
 XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
 PA  
 PI Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;  
 PI Puthalakath H, Strasser A;  
 XX  
 DR WPI: 1999-244030/20.  
 DR P-PSDR; AAW98158.  
 XX  
 PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer  
 treatment  
 XX  
 PS Claim 7: Page 101-102; 145pp; English.

CC The present sequence encodes the extra long form (EL) of human Bim,  
 or Bcl-2 interacting mediator of cell death (see AAW98158), a novel  
 member of the Bcl-2 family that is capable of inducing cell death  
 (apoptosis) and which acts as a 'death-ligand' for certain members  
 of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the  
 only Bcl-2 homology region which it encompasses is BH3. It is the  
 only BH3-only protein for which splice variants exist. These  
 result in the expression of a variety of isoforms, i.e. Bim-S,  
 Bim-L and Bim-EL. cDNAs encoding human Bim-L and Bim-EL (see  
 AAW98158) were isolated from embryo and liver cDNA libraries using  
 mouse b1m cDNA. Murine Bim-S, Bim-L and Bim-EL isoforms (see  
 AAW98154-56) are also provided. The human Bim gene maps to  
 chromosome 2 at bands 2q13-2q13. Binding the dynein light  
 chain was shown to regulate the pro-apoptotic activity of Bim.  
 Bim-S, the splice variant which does not bind to dynein light  
 chain, is a much more potent killer than either Bim-L or Bim-EL.  
 The invention provides variants (see AAW98159-68) of murine and human  
 Bim-L or Bim-EL that cannot bind, couple or otherwise associate  
 with a dynein light chain. The identification of Bim permits the  
 identification and rational design of a range of products for use  
 in therapy, diagnosis, antibody generation and involving modulation  
 of physiological cell death. These therapeutic molecules may act  
 as either antagonists or agonists of Bim's function and will be  
 useful in cancer, autoimmune or degenerative disease therapy.  
 Increased Bim expression or Bim activity is useful, e.g. for  
 treatment or prophylaxis in conditions such as cancer and deletion  
 of autoreactive lymphocytes in autoimmune disease. Decreased Bim



CC expression of Bim activity is useful in regulating inhibition or  
 CC prevention of cell death or degeneration such as under cytotoxic  
 CC conditions during e.g. gamma-irradiation and chemotherapy or during  
 CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,  
 CC hypoxia, degenerative diseases or for prolonging the survival of  
 CC cells being transplanted for treatment of disease. Since Bim is  
 CC expressed in germ cells, modulating Bim expression or Bim activity  
 CC is useful, e.g. as a contraceptive or method of sterilization by  
 CC preventing generation of fertile sperm.

XX Sequence 596 BP; 145 A; 175 C; 146 G; 130 T; 0 other;

alignment\_scores:  
 Quality: 1065.00 Length: 198  
 Ratio: 5.379 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-508-832-10 x AAX24997

Align seg 1/1 to: AAX24997 from: 1 to: 596

1 MetAlaLysGlnProSerAspValSerGluCysAspArgCluGlyAr 17  
 1 ATGCGAAGCAACCTTCTGATGTAAGTCTGAGTGTGACCGAGAGGTAG 50  
 17 gGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34  
 51 ACAATTCACGCTCGCGAGAGGCGCTCCCGAGCTCAGACCTGGGGCCCTA 100  
 34 hrSerLeuGlnThrGluProGlnGlnGlnGlnGlnGlnGlnGlnGlnG 50  
 101 CCTCCTTACACAGACAGACGACACAGTAACTCTGAAGGCAATCAGGAGGT 150  
 51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaP:pr 67  
 151 GAAGGGGACAGCTGCCCGACGCGAGCCCTCAGGGCGCTGGCCGCCACC 200  
 67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMet 84  
 201 TGCAGAGCCCTGGCCCTTTTGTACAGATCCCGCTTTTCATCTTATGA 250  
 84 rGArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 100  
 251 GAAGATCCTCCCTGCTCTCTCATCTCCAGTGGGTATTTCTCTTTTGAC 300  
 101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117  
 301 ACAGACAGAGAGCCACACCCATGAGTTGTGACAAATCAACACACACCCC 350  
 117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134  
 351 AAGTCTCTCTTGGCAGGCTTCAACCACTATCTCAGTGCATGGCTTCCA 400  
 134 eTArgGlnAlaGluProAlaAspMetArgProGluIleTTrpIleAlaGln 150  
 401 TGAGCGACGCTGAACCTGCAGATATCCCGCCAGACATATGATCGCCCAA 450  
 151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaArgVa 167  
 451 GAGTTCGGCGGTATCGGAGACGAGTTTAAACGCTTACTATGCAAGGAGGT 500  
 167 lPheLeuAsnAsnTyrGlnAlaAlaGluAspHisProArgMetValIleL 184  
 501 ATTTTGAATAATTAACAGACGCGGAGACACCCAGCAATGTTATCT 550  
 184 euArgLeuLeuArgTyrIleValArgLeuValTyrPheMetHis 198  
 551 TAGCACTGTACCTACATGTGCCCTTGTGGTGGAGAAATGCAT 594

seq\_name: /SIDSI/gc9data/geneseq/geneseqn-emb1/NA1999.DAT: AAX24995

seq\_documentation\_block:

ID AAX24995 standard; cDNA; 590 BP.

XX AAX24995:

XX 05-JUL-1999 (first entry)

XX Murine Bcl-2 interacting mediator of cell death Bim-EL cDNA.

XX Bim-EL; Bcl-2 interacting mediator of cell death; apoptosis;  
 KW cell cycle; mouse; cancer; autoimmune disease;  
 KW degenerative disease; therapy; contraceptive; splice variant;  
 KW isoform; ss.

XX Mus musculus.

XX W09914321-A1.

XX 25-MAR-1999.

XX 17-SEP-1998: 98WO-AU00772.

XX 24-SEP-1997: 97AU-0009373.

XX 17-SEP-1997: 97AU-0009263.

XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.

XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;  
 PI Puthalakath H, Strasser A;

XX WPI: 1999-244030/20.

XX P-PSDB: AAW98156.

XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer

XX treatment

XX Claim 3: Page 96-97: 145pp: English.

XX The present sequence encodes the extra long form (EL) of murine Bim,  
 CC or Bcl-2 interacting mediator of cell death (see AAW98156), a novel  
 CC member of the Bcl-2 family that is capable of inducing cell death  
 CC (apoptosis) and which acts as a 'death-ligand' for certain members  
 CC of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the  
 CC only Bcl-2 homology region which it encompasses is BH3. It is the  
 CC result in the expression of a variety of isoforms, i.e. Bim-S,  
 CC Bim-L and Bim-EL (see AAW98154-56). cDNAs encoding these murine Bim  
 CC isoforms were obtained from a T lymphoma cDNA library using human  
 CC recombinant Bcl-2 protein. The murine Bim gene has been mapped to  
 CC chromosome 2 at bands F3-G. Human Bim-L and Bim-EL isoforms have  
 CC also been identified (see AAW98157-59). Binding the dynein light  
 CC chain was shown to regulate the pro-apoptotic activity of Bim.  
 CC Bim-S, the splice variant which does not bind to dynein light  
 CC chain, is a much more potent killer than either Bim-L or Bim-EL.  
 CC The invention provides variants (see AAW98159-68) of murine and human  
 CC Bim-L or Bim-EL that cannot bind, couple or otherwise associate  
 CC with a dynein light chain. The identification of Bim permits the  
 CC identification and rational design of a range of products for use  
 CC in therapy, diagnosis, antibody generation and involving modulation  
 CC of physiological cell death. These therapeutic molecules may act  
 CC as either antagonists or agonists of Bim's function and will be  
 CC useful in cancer, autoimmune or degenerative disease therapy.  
 CC Increased Bim expression or Bim activity is useful, e.g. for  
 CC treatment of prophyllaxis in conditions such as cancer and deletion  
 CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim  
 CC expression of Bim activity is useful in regulating inhibition or  
 CC prevention of cell death or degeneration such as under cytotoxic  
 CC conditions during e.g. gamma-irradiation and chemotherapy or during  
 CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,  
 CC hypoxia, degenerative diseases or for prolonging the survival of  
 CC cells being transplanted for treatment of disease. Since Bim is  
 CC expressed in germ cells, modulating Bim expression or Bim activity  
 CC is useful, e.g. as a contraceptive or method of sterilization by



Percent Similarity: 69.697 Percent Identity: 69.697  
 alignment\_block:  
 US-09-508-832-10 x AAX24994  
 . Align seg 1/1 to: AAX24994 from: 1 to: 416

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17  
 1 AFGGCAAGCAACCTCTGATGTAAGTCTGAGTGTGACCCAGAGTAG 50  
 17 GcInLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34  
 51 ACAATTGCAGCTCGGAGAGGCTCCCGAGCTCAGACCTGGGCGCCCTA 100  
 34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50  
 101 COTCCCTACAGACAGCCCAAA..... 123  
 51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67  
 123 ..... 123  
 67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMetA 84  
 123 ..... 123  
 84 rGArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 100  
 123 ..... 123  
 101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117  
 124 ...GACAGGAGCCGAGCAGCCATGAGTGTGACAAATCAACACAAAGCCC 170  
 117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134  
 171 MAGTCCCTTGCAGGCTTCAACCACTATCTCAGTGCATGCGTTCOA 220  
 134 eArgGlnAlaGluProAlaAspMetArgProGluIleThrPheAlaGln 150  
 221 TGAGGAGGCTGACCCGAGATATCGCCGAGATATGATGATGATGATGAT 270  
 151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaArgTgVa 167  
 271 CAGTTGGCGGTATCGGAGACGAGTGTAAACGCTTACTATGCAACGAGGT 320  
 167 lPheLeuAsnAspTyrGlnAlaAlaGluAspHisProArgMetValIleL 184  
 321 ATTTTGTGATTAATACACAGCAGCCGAGACCCAGCAATGTTATCT 370  
 184 euArgLeuLeuArgTyrIleValArgLeuValTrpArgMetHis 198  
 371 TACGACTGTTACGTTACATTTGTCGCTGCGTGTGAGAGTGCAT 414

seq\_name: /SDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: AAX24994  
 seq\_documentation\_block:  
 ID\_AAX24994 standard; cDNA; 422 BP.  
 XX AC AAX24994;  
 XX DT 05-JUL-1999 (first entry)  
 XX DE Murine Bcl-2 interacting mediator of cell death Bim-L cDNA.  
 XX KW Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;  
 KW cell cycle; mouse; cancer; autoimmune disease;  
 KW degenerative disease; therapy; contraceptive; splice variant;  
 KW isoform; ss.  
 XX OS Mus musculus.  
 XX

PN WO9914321-A1.  
 XX 25-MAR-1999.  
 XX 17-SEP-1998; 98WO-AU00772.  
 XX 24-SEP-1997; 97AU-0009373.  
 PR 17-SEP-1997; 97AU-0009263.  
 XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
 XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;  
 PI puthalakath H, Strasser A;  
 XX P-PSDB; AAW98155.  
 DR WPI: 1999-244030/20.  
 XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer  
 PT treatment  
 PS Claim 3: Page 94-95; 145pp; English.  
 CC The present sequence encodes the long form (L) of murine Bim, or  
 CC Bcl-2 interacting mediator of cell death (see AAW98155), a novel  
 CC member of the Bcl-2 family that is capable of inducing cell death  
 CC (apoptosis) and which acts as a 'death-ligand' for certain members  
 CC of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the  
 CC only Bcl-2 homology region which it encompasses is BH3. It is the  
 CC only BH3-only protein for which splice variants exist. These  
 CC result in the expression of a variety of isoforms, i.e. Bim-S,  
 CC Bim-L and Bim-EL (see AAW98154-56). cDNAs encoding these murine Bim  
 CC isoforms were obtained from a T lymphoma cDNA library using human  
 CC recombinant Bcl-2 protein. The murine Bim gene has been mapped to  
 CC chromosome 2 at bands F3-G. Human Bim-L and Bim-EL isoforms have  
 CC also been identified (see AAW98157-58). Binding the dynein light  
 CC chain was shown to regulate the pro-apoptotic activity of Bim.  
 CC Bim-S, the splice variant which does not bind to dynein light  
 CC chain, is a much more potent killer than either Bim-L or Bim-EL.  
 CC The invention provides variants (see AAW98159-68) of murine and human  
 CC Bim-L or Bim-EL that cannot bind, couple or otherwise associate  
 CC with a dynein light chain. The identification of Bim permits the  
 CC identification and rational design of a range of products for use  
 CC in therapy, diagnosis, antibody generation and involving modulation  
 CC of physiological cell death. These therapeutic molecules may act  
 CC as either antagonists or agonists of Bim's function and will be  
 CC useful in cancer, autoimmune or degenerative disease therapy.  
 CC Increased Bim expression or Bim activity is useful, e.g. for  
 CC treatment or prophylaxis in conditions such as cancer and deletion  
 CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim  
 CC expression or Bim activity is useful in regulating inhibition or  
 CC prevention of cell death or degeneration such as under cytotoxic  
 CC conditions during e.g. gamma-irradiation and chemotherapy or during  
 CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,  
 CC hypoxia, degenerative diseases or for prolonging the survival of  
 CC cells being transplanted for treatment of disease. Since Bim is  
 CC expressed in germ cells, modulating Bim expression or Bim activity  
 CC is useful, e.g. as a contraceptive or method of sterilization by  
 CC preventing generation of fertile sperm.  
 XX SQ Sequence 422 BP; 112 A; 116 C; 109 G; 85 T; 0 other;

alignment\_scores:  
 Quality: 589.00 Length: 200  
 Ratio: 4.566 Gaps: 2  
 Percent Similarity: 64.500 Percent Identity: 60.000  
 alignment\_block:  
 US-09-508-832-10 x AAX24994  
 Align seg 1/1 to: AAX24994 from: 1 to: 422

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17

```
|||||
1 ATGGCCAGACACCTTCGATGTAGTTCTGAGTGTGACAGAGAGGTGG 50
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34
51 ACAATTCAGGCTCTCTGAGAGGCTCCCGAGCTCAGGCTGGGGCCCTA 100
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyGly 50
101 CCTCCTACAGACAGAACCGCAA..... 123
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
123 ..... 123
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMetA 84
123 ..... 123
84 rArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheAsp 100
123 ..... 123
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
124 ...GACAGAGCCGGCCACCAGTGTGTGACAGTCAACACAAACCC 170
117 oSerProProCysGlnAlaPheAsnHisLysLeuSerAlaMetAlaSerM 134
171 AAGTCCTCTTCCAGGCTTCAACCACTATCTCAGTGCATGCAATGCTTCCA 220
134 eTargGlnAla.....GluProAlaAspMetArgProGluIleTrpIle 148
221 TAGCAGCTCTCAGAGGACCTGAGTGTGCTGCGCCCGGAGATACGGATT 270
149 AlaGlnGluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyraLar 165
271 GCACAGAGAGCTGCGCGGATCGGAGACGAGTTCACGCAAACTTACACAG 320
165 gATGValPheLeuAsnAsnTyrGlnAlaAlaGluAspHisProArgMetV 182
321 GAGGCGTGTTCGAATGATTACCGCGAGGCTGAAGACCAACCCCAATGG 370
182 alrLeuArgLeuLeuArgTyrIleValArgLeuValTrpArgMetHis 198
371 TTATCTTACAACTGTTCCTTTATCTCTCTGTTGAGAGAGGCAT 420
seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA565350
seq_documentation_block:
ID AA565350 standard; cDNA; 442 BP.
XX
AC AA565350;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #1154.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
```

Drmanac RT, Liu C, Tang YT;  
WPI: 2001-639362/73.  
P-PSDB: AB601163.  
New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity  
Claim 1: SEQ ID No 1154; 103pp; English.  
The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.  
Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
Sequence 442 BP; 94 A; 154 C; 105 G; 89 T; 0 other;

alignment\_scores:  
Quality: 497.00 Length: 92  
Ratio: 5.402 Gaps: 0  
Percent similarity: 100.000 Percent identity: 100.000

alignment\_block:  
US-09-508-832-10 x AAS65350 ..  
Align seg 1/1 to: AAS65350 from: 1 to: 442

1 MetAlaLysGlnProSerAspValSerSerCysAspArgGluGlyAr 17  
167 ATGGCAAGCAACCTTCTGATGTAAAGTCTGAGTGTGACGAGAGGTAG 216  
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34  
217 ACAATTCAGGCTCTCTGAGAGGCTCCCGAGCTCAGGCTGGGGCCCTA 266  
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyGlyGly 50  
267 CCTCCTACAGACAGACAGCCACAGGTAATCTCTGAGGCAATCAGCGAGGT 316  
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67  
317 GAAGGGGACAGCTGCCCCACGCGAGCCCTCAGGGCCCGCTGGCCCCACC 366  
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMetA 84  
367 TGCCAGGCCCTGGGCCCTTTGCTACACAGATCCCGCTTTTCATCTTTATGA 416  
84 rArgSerSerLeuLeuSerArgSer 92  
417 GAAGATCCTCCTGCTGCTCGCATCC 442

seq\_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AA524993

seq\_documentation\_block:  
ID\_AAX24993 standard; cDNA; 332 BP.  
XX  
AC AAX24993;  
XX  
DT 05-JUL-1999 (first entry)  
XX  
DE Murine Bcl-2 interacting mediator of cell death Bim-S cDNA.  
XX  
KW Bim-S; Bcl-2 interacting mediator of cell death; apoptosis;  
KW cell cycle; mouse; cancer; autoimmune disease;  
KW degenerative disease; therapy; contraceptive; splice variant;  
KW isoform; SS.  
XX  
OS Mus musculus.  
XX  
XX WO9914321-A1.  
XX  
XX 25-MAR-1999.  
XX  
XX 17-SEP-1998; 98WO-A000772.  
XX  
XX 24-SEP-1997; 97AU-0009373.  
XX  
XX 17-SEP-1997; 97AU-0009263.  
XX  
XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
XX  
XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;  
XX Puthalakath H, Strasser A;  
XX  
XX WPI; 1999-244030/20.  
XX  
XX P-PSDB; AAW98154.  
XX  
XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer  
XX treatment  
XX  
XX Claim 3; Page 92; 145pp: English.  
XX  
XX The present sequence encodes the short form (S) of murine Bim, or  
XX Bcl-2 interacting mediator of cell death (see AAW98154), a novel  
XX member of the Bcl-2 family that is capable of inducing cell death  
XX (apoptosis) and which acts as a 'death-ligand' for certain members  
XX of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the  
XX only Bcl-2 homology region which it encompasses is BH3. It is the  
XX only BH3-only protein for which splice variants exist. These  
XX result in the expression of a variety of isoforms, i.e. Bim-S,  
XX Bim-L and Bim-EL (see AAW98154-56). cDNAs encoding these murine Bim  
XX isoforms were obtained from a T lymphoma cDNA library using human  
XX recombinant Bcl-2 protein. The murine Bim gene has been mapped to  
XX chromosome 2 at bands F3-C. Human Bim-L and Bim-EL isoforms have  
XX also been identified (see AAW98157-58). Binding the dynein light  
XX chain was shown to regulate the pro-apoptotic activity of Bim.  
XX Bim-S, the splice variant which does not bind to dynein light  
XX chain, is a much more potent killer than either Bim-L or Bim-EL.  
XX The invention provides variants (see AAW98159-68) of murine and human  
XX Bim-L or Bim-EL that cannot bind, couple or otherwise associate  
XX with a dynein light chain. The identification of Bim permits the  
XX identification and rational design of a range of products for use  
XX in therapy, diagnosis, antibody generation and involving modulation  
XX of physiological cell death. These therapeutic molecules may act  
XX as either antagonists or agonists of Bim's function and will be  
XX useful in cancer, autoimmune or degenerative disease therapy.  
XX Increased Bim expression or Bim activity is useful, e.g. for  
XX treatment or prophylaxis in conditions such as cancer and deletion  
XX of autoreactive lymphocytes in autoimmune disease. Decreased Bim  
XX expression of Bim activity is useful in regulating inhibition or  
XX prevention of cell death or degeneration such as under cytotoxic  
XX conditions during e.g. gamma-irradiation and chemotherapy or during  
XX HIV/AIDS or other viral infections, ischemia, myocardial infarction,  
XX hypoxia, degenerative diseases or for prolonging the survival of  
XX cells being transplanted for treatment of disease. Since Bim is  
XX expressed in germ cells, modulating Bim expression or Bim activity

is useful, e.g. as a contraceptive or method of sterilization by  
preventing generation of fertile sperm.  
XX  
SQ Sequence 332 BP; 87 A; 85 C; 91 G; 69 T; 0 other;  
  
alignment\_scores:  
Quality: 406.00 Length: 200  
Ratio: 4.101 Gaps: 2  
Percent Similarity: 49.500 Percent Identity: 45.000  
  
alignment\_block:  
US-09-508-832-10 x AAX24993 ..  
  
Align seg 1/1 to: AAX24993 from: 1 to: 332  
  
1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGly 17  
|||||  
1 ATGGCCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGTGG 50  
  
17 GGLnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34  
|||||  
51 ACAATTGGAGGCTCTGAGAGGCTCCCAAGCTCAGGCTGGGGCCCTA 100  
  
34 hSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGly 50  
|||||  
101 CCTCCCTACAGACAGAACCCGCA.. 123  
  
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPro 67  
123 ..... 123  
  
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPhePheMeta 84  
123 ..... 123  
  
84 rGArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheAsp 100  
123 ..... 123  
  
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117  
123 ..... 123  
  
117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134  
124 .....GCTTCCA 130  
  
134 eArgGlnAla.....GluProAlaAspMetArgProGluLeuTrpIle 148  
:::|||||: ||||| :|||:|||||: |||  
131 TACGACAGCTCTCAGCAGCAACCTGAGATCTGCCCGCGAGATACGATT 180  
  
149 AlaGlnGluLeuArgGlyLeuGlyAspGluPheAsnAlaTyrTyrAla 165  
|||||  
181 GCACAGGAGCTCGCGCGGATCGGAGACAGCTTCACGAACTTACACA 230  
  
165 gArgValPheLeuAsnAsnTyrGlnAlaAlaGluAspHisProArgMetV 182  
:::|||||: ||||| :|||:|||||: |||  
231 GAGGTGTTTGCAAATGATTACCGGAGGCTGAAGACACCTCAATGG 280  
  
182 alileuArgLeuLeuArgTyrIleValArgLeuValTrpArgMethis 198  
|||||  
281 TTATCTTAACTGTACGCTTTATCTTCGCTGTGTATGGAGAAGCAT 330  
  
seq\_name: /SIDSI/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:RAI89284  
seq\_documentation\_block:  
ID AAI89284 standard; cDNA; 371 BP.  
XX  
AC AAI89284;  
XX  
DT 06-NOV-2001 (first entry)  
XX

DE Human polynucleotide SEQ ID NO 9344.  
 XX  
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; Cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164835-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 XX 26-FEB-2001; 2001WO-US04927.  
 PF  
 XX 28-FEB-2000; 2000US-0515126.  
 PR  
 PR 18-MAY-2000; 2000US-0577409.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Tang YT, Liu C, Drmanac RT;  
 PI  
 XX WPI; 2001-514838/56.  
 XX  
 DR P-PSDB; AA009353.  
 DR  
 XX Isolated nucleic acids and polypeptides, useful for preventing  
 PT diagnosing and treating e.g. leukaemia, inflammation and immune  
 PT disorders -  
 PT  
 XX  
 PS Claim 1; SEQ ID NO 9344; 1399pp + Sequence Listing; English.  
 XX  
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 371 BP; 106 A; 95 C; 81 G; 89 T; 0 other;

alignment\_scores:  
 Quality: 123.00 Length: 23  
 Ratio: 5.348 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-508-832-10 x AAI89284/rev ..  
 Align seg 1/1 to reverse of: AAI89284 from: 1 to: 371  
 176 GluAspHisProArgMetValIleLeuArgLeuArgTyrIleValAr 192  
 371 GAAGACACACCCAGCAAGTGTATCTACGACTGTTACGTTACATGTCGC 322

192 gLeuValTrpArgMethHis 198  
 321 CTGTGTGTGGAGATGCAT 303

seq\_name: /SIDSL/ycgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV35620

seq\_documentation\_block:

ID AAV35620 standard; DNA; 32367 BP.  
 XX  
 AC AAV35620;

XX  
 DT 07-SEP-1998 (first entry)  
 XX  
 DE Human SHOX (short stature homeobox containing gene) gene sequence.  
 KW Homeobox domain; human growth gene; growth regulation; growth defect;  
 KW Turner's syndrome; short stature homeobox containing gene; SHOXa;  
 KW SHOX; bone disease; osteoporosis; calcium regulation; short stature;  
 KW transcription factor A; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9814568-A1.  
 XX  
 PD 09-APR-1998.  
 XX  
 XX 29-SEP-1997; 97WO-EP05355.  
 PF  
 XX 16-JAN-1997; 97EP-0100583.  
 PR  
 PR 01-OCT-1996; 96US-0027633.  
 XX  
 XX (RAPP/) RAPPOLD-HOERBRAND G.  
 PA  
 XX Rao E, Rappold-hoerbrand G;  
 PI  
 XX WPI; 1998-271719/24.  
 DR  
 XX New human growth genes - used to develop products for the diagnosis  
 PT and treatment of human growth defects such as short stature, e.g.  
 PT Turner's syndrome  
 PT  
 XX  
 PS Claim 19; Pages 51-67; 84pp; English.  
 XX  
 CC This is the human SHOX gene sequence containing the PAR1 region. The  
 CC gene region corresponding to short stature has been identified as a  
 CC region of approximately 500 kb in the PAR1 region of the X and Y  
 CC chromosomes. Three genes in this region have been identified as  
 CC candidates for the short stature gene. These genes were designated SHOX  
 CC (also referred to as SHOX93 or HOX93), PET92 and SHOT (SHOX-like homeobox  
 CC gene on chromosome three). The SHOX gene has two separate splicing sites  
 CC resulting in two variations SHOXa and SHOXb. The specification provides  
 CC sequences of SHOX (short stature homeobox-containing) genes SHOX ET92,  
 CC SHOXa, SHOXb, SHOT and exons of the SHOX genes as shown in AAV35610 to  
 CC AAV35621 and protein sequences of the human growth protein transcription  
 CC factor SHOXa, SHOXb and SHOT as shown AAW60573 to AAW60575. The novel  
 CC genes are responsible for human growth. Defects in the genes can cause  
 CC short stature, e.g. Turner's syndrome. The products can be used to  
 CC develop agents for the treatment of short stature or other human growth  
 CC disorders. The products can also be used for providing a mitogenic effect  
 CC on cells, e.g. for the treatment of bone diseases such as osteoporosis  
 CC and diseases involved with disturbance in the bone calcium regulation.  
 XX  
 SQ Sequence 32367 BP; 7627 A; 8130 C; 8564 G; 8043 T; 3 other;

alignment\_scores:  
 Quality: 119.50 Length: 177  
 Ratio: 1.271 Gaps: 13  
 Percent Similarity: 53.107 Percent Identity: 33.898

alignment\_block:  
 US-09-508-832-10 x AAV35620/rev ..

Align seg 1/1 to reverse of: AAV35620 from: 1 to: 32367  
 9 SerSerGluCysAspArgGluGlyArg.....GlnLeuGlnProAlaG1 23  
 14144 TCCCAAGGCTCTCCCGGAGACACACGAGTGTCCCAAGGCTCTCCCGGGTCCAC 14095  
 23 uArgProProGlnLeu.....ArgProGlyAlaProThrSerLeuG 37  
 14094 CAGGTCCCAAGGCTCTCTCCCGAGACACACGAGTGTCCCGGCTCTCCGC 14045











1749 ACCGTCCATGAACCAACCTGCCTCAGTGGATGCCAACCAACCAAGCTGA 1700

116 hrProSerProCysGlnAlaPhe...AsnHis 126

1699 CGCGGTCACTGCCTGCCAGGCTCAGTGATCAT 1665

seq\_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAA58472

seq\_documentation\_block:

ID AAA58472 standard: DNA; 18660 BP.

AC AAA58472;

31-OCT-2000 (first entry)

Nucleotide sequence of the bleomycin (BLM) gene cluster ORFs 31-40.

BLM gene cluster; bleomycin gene cluster; polyketide metabolite; bleomycin; bleomycin analogue; holo-carrier protein; thiazolidine; thiazoline; bithiazoline; microbial metabolite; sugar; ss.

Streptomyces verticillus.

Key Location/Qualifiers  
CDS 1..354  
/\*tag= a  
/transl\_except= (pos: 1..3, aa: Met)  
/note= "ORF31; encodes AAB07580"  
CDS 925..9450  
/\*tag= b  
/note= "ORF32; encodes AAB07581"  
CDS 9447..10802  
/\*tag= c  
/note= "ORF33; encodes AAB07582"  
CDS 10864..11877  
/\*tag= d  
/note= "ORF34; encodes AAB07583"  
CDS 11888..12640  
/\*tag= e  
/note= "ORF35; encodes AAB07584"  
CDS 12637..13920  
/\*tag= f  
/note= "ORF36; encodes AAB07585"  
CDS 13928..14746  
/\*tag= g  
/transl\_except= (pos: 1..3, aa: Met)  
/note= "ORF37; encodes AAB07586"  
CDS 14743..16479  
/\*tag= h  
/note= "ORF38; encodes AAB07587"  
CDS 16476..17462  
/\*tag= i  
/transl\_except= (pos: 1..3, aa: Met)  
/note= "ORF39; encodes AAB07588"  
CDS 17646..18659  
/\*tag= j  
/note= "ORF40; encodes AAB07589"  
/note= "no termination codon given"

WO200040704-A1.

13-JUL-2000.

06-JAN-2000; 2000WO-US00445.

06-JAN-1999; 99US-0115435.

05-FEB-1999; 99US-0118848.

05-JAN-2000; 2000US-0477962.

(RECC ) UNIV CALIFORNIA.

Shen B, Du L, Sanchez C, Chen M, Edwards DJ;

DR WPI: 2000-465974/40.  
DR P-PSDB: AAB07580, AAB07581, AAB07582, AAB07583, AAB07584, AAB07585.  
DR AAB07586, AAB07587, AAB07588, AAB07589.  
XX  
PT New bleomycin gene cluster components useful for peptide and/or  
PT polyketide metabolites, especially bleomycin, production and for  
PT chemically modifying biological molecules -  
XX  
PS Claim 8: Page 137-153; 162pp; English.  
XX  
CC The present sequence represents the BLM (bleomycin) gene cluster,  
CC containing open reading frames (ORFs) 31-40. The proteins encoded  
CC by the gene cluster are useful for producing peptides and/or polyketide  
CC metabolites, especially bleomycin or bleomycin analogues. They are  
CC also useful for chemically modifying biological molecules to produce  
CC branched methyl groups, and for coupling amino acids and fatty  
CC acids. They may be reacted with an apo-carrier protein and coenzyme A  
CC to produce a holo-carrier protein. The BLM gene cluster or catalytic  
CC domains can be used individually or collectively to produce  
CC thiazolidine, thiazoline, bithiazoline and bithiazoline-containing  
CC microbial metabolites. The BLM gene cluster may also be used to produce  
CC sugars.  
XX  
S0 Sequence 18660 BP; 2365 A; 7239 C; 6510 G; 2546 T; 0 other;

alignment\_scores:  
Quality: 112.00 Length: 192  
Ratio: 1.333 Gaps: 9  
Percent Similarity: 43.750 Percent Identity: 27.604  
alignment\_block:  
US-09-508-832-10 x AAA58472/rev  
Align seg 1/1 to reverse of: AAA58472 from: 1 to: 18660  
14 ArgGluGlyArgGlnLeuGlnPro.....AlaGluArgPr 25  
6802 CGACGCGCGCGAGGCGCCACTCCGGATCCACTGGCAGCATAGACGCCGCC 6753  
25 oProGlnLeuArgProGlyAlaProThrSerLeuGlnThrGluProGlnG 42  
6752 CCCCCAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6703  
42 LysAsnProGlu.....GlyAsn 47  
6702 GCACAGCGCGAGCGCGCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6653  
48 HisGlyGlyGlyGlyAspSerCysPro.....HisGlyLys 59  
6652 CGGCAGCGCGCGGTGACCGCGGTGTCCAGCTCGCGGAGCTCAGCGCTCGAG 6603  
59 rProGlnGlyProLeu...AlaProProAlaSerProGlyProPheAlaT 75  
6602 TCCCGGTGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6553  
75 hrArg..... 76  
6552 GAAGCGGTACGGAGCGGTCTCCGCCCGCGCGCGCGCGCGCGCGCGCGCG 6503  
77 .....SerProLeuPheIlePheMetArgArgSerSerLeuLeuSerAr 91  
6502 CGGACCG 6465  
91 gSerSerSerGlyTyPheSerPheAspThrAspArgSerProAlaProm 108  
6464 CTCAGCGCGAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6418  
108 etSerCysAspLysSerThrGlnThrProSerProProCysGlnAlaPhe 124  
6417 GCCGAGGAGGTCAGGTAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 6368  
125 AsnHisTyrLeuSerAlaMetAlaSerMetArgGlnAlaGluProAlaAs 141

[illegible]

```

6367 CGAACAGGTGGGACAGT...:::||||| |||| ||||||
141 pMetArgProGlutIleTrpIleAlaGlnGluteuArgArgIleClyAspC 158
6320 CCGAGA.....||||| |||
158 luPheAsnAlaTyrtYrAlaArGArg 166
6297 GTCGAACCGGTGTCTCACGCCCGCA 6272

seq_name: /SIDS1/gcgdata/genesesq/geneseqn-emb1/NA2001A.DAT-AAI66848
seq_documentation_block:
ID AAI66848 standard; DNA; 3090 BP.
XX AC AA166848;
XX AC
XX AC
DT 07-JAN-2002 (first entry)
DE Nucleotide sequence of HSPDE4AAcatD-EGFP fusion.
DE DE
KW PDE4: central nervous system; antiinflammatory; cytostatic; nootropic;
KW autoimmune; ischemic; osteopathic; GFP; green fluorescent protein;
KW fusion protein; ss.
XX Homo sapiens.
XX OS Aequorea victoria.
XX OS
XX FH Location/Qualifiers
FT CDS 1..3090
FT FT /note= "fusion protein"
FT PN WO200179536-A2.
FT PD
FT PP 11-APR-2001; 2001WO-DK00264.
XX PR
XX PR 17-APR-2000; 2000DK-0000651.
XX PR 29-NAY-2000; 2000DK-0000849.
XX PA (BIOI-) BIOIMAGE AS.
XX PI Terry BR, Scudder KM, Bjorn SP, Thastrup O, Almholzt DC;
XX PI Praestegaard M;
XX XX
XX WPI: 2001-611727/70.
XX DR P-PSDB; AAG65779.
XX PT
XX PT Determining if a compound is a dislocator of PDE4 for identifying
XX compounds for treating CNS and inflammatory disease comprises
XX identifying compounds which remove PDE4 spots -
XX XX
XX Example 1; Page 142-147; 160pp; English.
XX PS
XX PS The invention relates to determining, if a compound, is a dislocator of
XX PDE4. The method comprises testing if the compound removes PDE4 spots,
XX which may optionally be induced by a Rolipram-like reference compound,
XX and testing if it inhibits the catalytic activity of the PDE4, where the
XX compound is a dislocator of PDE4, if it removes PDE spots and if it does
XX not inhibit the catalytic activity of PDE4. The method is useful for
XX identifying compounds useful for the treatment of diseases of the central
XX nervous system such as depression and for the treatment of inflammatory
XX disease such as joint inflammation, Crohn's disease, inflammatory bowel
XX disease, respiratory infections, chronic obstructive pulmonary disease
XX (COPD), including asthma, chronic bronchitis, pulmonary emphysema,
XX endotoxias shock, toxic shock syndrome, systemic lupus erythematosia,
XX psoriasis, bone resorption diseases, reperfusion injury, cancer and HIV
XX infection. The use of a reagent that can mimic or reverse the effect of
XX the compound with affinity for the catalytic site on intracellular
XX distribution of the PDE for the preparation of a medicament. The present
XX sequence represents the nucleotide sequence of a HSPDE4AAcatD-EGFP

```

OM of: US-09-508-832-10 to: Issued\_Patents\_NA:\* out\_format : pfs

Date: Aug 7, 2002 10:36 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framer-p2n.spool -DEV=xlh  
-O7/cgn2\_1/USPTO\_model/US09508832/runat\_06082002\_103611\_2799/app\_query.fasta.1.259  
-DB=Issued\_Patents\_NA -QWTF=fastap -SUFFIX=rni -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-XGAPOP=6.000 -XGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsum62  
-TRANS=humana40.cdi -LIST=45 -DOCLIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZE=500 -NINLEN=0 -MAXLEN=2000000000  
-USBR=US09508832.ecgn1\_1.46 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLIFY -WAIT -THREADS=1

Search information block:

Query: US-09-508-832-10

Query length: 198

Database: Issued\_Patents\_NA.\*

Database sequences: 383533

Database length: 122816752

Search time (sec): 46.200000

score list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-474-379C-64	111.50	147.06	1.67	3705	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-146-249A-64	111.50	147.06	1.67	3705	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-206-188B-64	111.50	147.06	1.67	3705	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-997-080-115	107.00	158.63	0.3796	647	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-997-362-115	107.00	158.63	0.3796	647	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-095-855-115	107.00	158.63	0.3796	647	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-324-542-115	107.00	158.63	0.3796	647	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-602-359A-24	107.00	153.54	0.7291	1041	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-997-080-113	107.00	149.15	1.28	1569	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-997-362-113	107.00	149.15	1.28	1569	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-095-855-113	107.00	149.15	1.28	1569	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-324-542-113	107.00	149.15	1.28	1569	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-997-080-159	107.00	148.77	1.34	1626	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-997-362-159	107.00	148.77	1.34	1626	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-095-855-159	107.00	148.77	1.34	1626	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-324-542-159	107.00	148.77	1.34	1626	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-144-085-3	106.50	115.60	94.63	33529	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-105-537-9	106.00	148.35	1.42	1458	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-320-878-24	106.00	147.60	1.56	1565	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-135-994-2	106.00	145.73	1.99	1864	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-105-537-3	106.00	124.45	30.40	13613	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-998-416-1077	104.50	156.02	0.5306	571	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-484-840-1	104.00	143.09	2.79	1776	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-483-094-1	104.00	143.09	2.79	1776	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-09-318-448-13	104.00	143.09	2.79	1776	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-103-840A-2	104.00	59.46	1.2e+05	4403765	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-09-103-840A-1	104.00	59.44	1.2e+05	4411529	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-764-233A-1	103.00	105.93	326.99	49377	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-09-335-409-1	102.50	101.60	569.85	68750	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-568-102-1	102.50	101.60	569.85	68750	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-567-969-1	102.50	101.60	569.85	68750	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-568-480-1	102.50	101.60	569.85	68750	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-568-486-1	102.50	101.60	569.85	68750	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-568-472-1	102.50	101.60	569.85	68750	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-443-501A-2	102.00	100.32	671.70	71989	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-231-193A-44	101.50	132.70	10.56	3243	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-486-273A-44	101.50	132.70	10.56	3243	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-480-474-44	101.50	132.70	10.56	3243	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-940-086A-44	101.50	132.70	10.56	3243	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-940-035A-44	101.50	132.70	10.56	3243	

/cgn2\_6/ptodata/1/ina/5B\_COMB.seq:US-08-231-193A-43 + 101.50 131.29 12.65  
/cgn2\_6/ptodata/1/ina/5B\_COMB.seq:US-08-486-273A-43 + 101.50 131.29 12.65  
/cgn2\_6/ptodata/1/ina/6A\_COMB.seq:US-08-480-474-43 + 101.50 131.29 12.65  
/cgn2\_6/ptodata/1/ina/6A\_COMB.seq:US-08-940-086A-43 + 101.50 131.29 12.65  
/cgn2\_6/ptodata/1/ina/6B\_COMB.seq:US-08-940-035A-43 + 101.50 131.29 12.65

seq\_name: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:US-08-474-379C-64

seq\_documentation\_block:  
: Sequence 64, Application US/08474379C  
: Patent No. 5977305  
: GENERAL INFORMATION:  
: APPLICANT: Wigler, Michael H.  
: APPLICANT: Colicelli, John J.  
: TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED  
: TITLE OF INVENTION: PROCESSES  
: NUMBER OF SEQUENCES: 88  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
: STREET: 233 South Wacker Drive/6300 Sears Tower  
: CITY: Chicago  
: STATE: Illinois  
: COUNTRY: United States of America  
: ZIP: 60606-6402  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/474,379C  
: FILING DATE: 07-JUN-1995  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US-07/511,715  
: FILING DATE: 20-APR-1990  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/206,188  
: FILING DATE: 01-MAR-1994  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 07/688,352  
: FILING DATE: 19-APR-1991  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Clough, David W.  
: REGISTRATION NUMBER: 36,107  
: REFERENCE/DOCKET NUMBER: 27866/32771  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (312) 474-6300  
: TELEFAX: (312) 474-0448  
: INFORMATION FOR SEQ ID NO: 64:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 3705 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: CDNA  
: FEATURE: CDS  
: NAME/KEY: CDS  
: LOCATION: 116..2773  
: US-08-474-379C-64

alignment\_scores:  
Quality: 111.50 Length: 171  
Ratio: 1.282 Gaps: 7  
Percent Similarity: 50.877 Percent Identity: 27.485

alignment\_block:  
US-09-508-832-10 x US-08-474-379C-64

Align seg 1/1 to: US-08-474-379C-64 from: 1 to: 3705

6 SerAspValSerSerGluCysArgGluGlyArgGlnLeuGlnProAl 22



```

seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:US-08-206-188B-64
seq_documentation_block:
: Sequence 64, Application US/08206188B
: Patent No. 6100025
: GENERAL INFORMATION:
: APPLICANT: Wlgler, Michael H.
: APPLICANT: Colicelli, John J.
: TITLE OF INVENTION: Cloning by Complementation and Related
: NUMBER OF SEQUENCES: 84
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Cerstein, Murray & Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606-6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/206.188B
: FILING DATE: 01-MAR-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/511,715
: FILING DATE: 20-APR-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: Clough, David W.
: REGISTRATION NUMBER: 36107
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312/474-6300
: TELEFAX: 312/474-0448
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 64:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3705 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 116..2774
: US-08-206-188B-64

alignment_scores:
Quality: 111.50 Length: 171
Ratio: 1.282 Gaps: 7
Percent Similarity: 50.877 Percent Identity: 27.485

alignment_block:
US-09-508-832-10 x US-08-206-188B-64 ..
Align seg 1/1 to: US-08-206-188B-64 from: 1 to: 3705

6 SerAspValSerGluCysAspArgGluGlyArgGlnLeuGlnProAl 22
||||| .....:||||| :
260 TCCGACAGCGCGAGCGCGCGGAGCGGCGCGCGCGCGCCAT 309
||||| .....:||||| :
22 aGluArgProProGlnLeu.....ArgProGlyAlaProThrS 35
||||| .....:||||| :
310 AGAGCGCGCGCGATGCGTGGACACAGCGCGCGCGCGCGCGCG 359
||||| .....:||||| :
35 erLeuGlnThrGluProGln.....GlyAsnProGluGly 46
:: .....:||||| :
360 CCGCGCATGTCCTGGCCCTCGTCTTCATGGCACTGGCGCGCGCG 409

```

```

47 AsnHisGlyGly.....GluGlyAspSerCysProH1 57
::: ||||| .....:||||| :
410 GGCGCGCGCGAGCGAGCAGCGCGCTTCGAGCGCAGAGTGGCGCGAC 459
57 sGlySerPro...GlnGlyProLeuAlaProProAlaSerProGlyProP 73
||||| .....:||||| :
460 ACCATCTCTGCGCGAGCGCGCTGGACTCGCAGCGGAGCGCGAGGA...C 506
::: .....:||||| :
73 heAlaThrArgSerProLeuPheLePheMetArgArgSerSerLeuLeu 89
::: .....:||||| :
507 TCGTGCTGCACGCCGGGGCGGCCACCGCCAGCGCGGAGTCTCTTCCTG 556
||||| .....:||||| :
90 SerArgSerSerSerGlyTyPheSerPheAspThrAspArgSerProAl 106
||||| .....:||||| :
557 TACCGCTCAGACAGCGACTAT.....GACATGTACCCCAA 591
::: .....:||||| :
106 aProMetSerCysAspLysSerThrGlnThrProSerProProCysGlnA 123
||||| .....:||||| :
592 GACCATGTCCCGGAC..... 607
123 laPheAsnHisTyrLeuSerAlaMetAlaSerMetArgGlnAlaGluPro 139
||||| .....:||||| :
608 .....TCTCGGTCCACGAGCGCGCGCGCTGAGAGC 640
140 AlaAspMetArgProGluLeuTrpIleAlaGlnGluLeuArgArgIleG1 156
::: ||| .....:||||| :
641 CTCATCGTAAACACCATTTCTCAGGTGCTGCCAGCGCTCGGAGCGCTCG 690
::: .....:||||| :
156 yAspGluPheAsn 160
::: .....:||||| :
691 TAGCAACTCTCTCA 703

seq_name: /cgn2_6/ptodata/1/lna/5B_COMB.seq:US-08-997-080-115
seq_documentation_block:
: Sequence 115, Application US/08997080
: Patent No. 5968524
: GENERAL INFORMATION:
: APPLICANT: NATSON, JAMES D.
: APPLICANT: TAN, PAUL L.J.
: TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
: NUMBER OF SEQUENCES: 194
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lev Offices of Ann W. Speckman
: STREET: 2601 Elliott Avenue, Suite 4185
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98121
: COMPUTER READABLE FORM:
: COMPUTER: IBM Compatible
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/997,080
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Sleath, Janet
: REGISTRATION NUMBER: 37,007
: REFERENCE/DOCKET NUMBER: 11000.1007
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 206-269-0565
: TELEFAX: 206-269-0563
: TELEX:
: INFORMATION FOR SEQ ID NO: 115:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 647 base pairs
: TYPE: nucleic acid

```

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic RNA  
US-08-997-080-115

## alignment\_scores:

Quality: 107.00 Length: 162  
Ratio: 1.259 Gaps: 9  
Percent Similarity: 52.469 Percent Identity: 30.864

## alignment\_block:

US-09-508-832-10 x US-08-997-080-115

Align seg 1/1 to: US-08-997-080-115 from: 1 to: 647

```
17 ArgGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPr 33
|||||
8 AGACAATTGGGTATGACGAAGAGCGCGCGTGGCTCGAGCGGGGCTC 57
|||||
33 OthrSerLeuGlnThrGluProGlnGlyAsnPro GluGlyAsnHisGly 49
|||||
58 AACGCCCTCGCAGAGC .....CCGTAAAGGTGACGTTGGG 92
|||||
50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
|||||
93 CCCGAGCGCTCGCAGACCTGCTGCTGGAGAAAGTGGGGCG .....CC 135
|||||
66 oProAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83
|||||
136 CCCAGCATCACCA...ACGATGGTGTGTCCATCGCCA..... 169
|||||
83 etArgArgSerLeuLeuSerArgSerSerSerGlyTyrPheSerPhe 99
|||||
170 ...AGGATCGAGCTGGAGGACCGCTACGAGAGATCGCGCTGAGCTGG 217
|||||
100 AspThrAspArgSerProAlaPrometSerCysAspLysSerThrGlnTh 116
|||||
218 ...TCAAAGAGTGGCGAGAGACGACGACGTCGCGGGGCGGCGCACC 264
|||||
116 rProSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlas 133
|||||
265 ACCACCGCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 314
|||||
133 erMetArgGlnAlaGluProAlaAspMetArgProGluIleTrrp..... 147
|||||
315 CGTCCGACCGCGCCACACCGCT .....CGGCTCAAGCTGCGCATCG 358
|||||
148 .....lleAlaGlnGluLeuArgArgleg 156
|||||
359 AGAAGGCTGTGAGGCTGTACCCAGCTGCTGCTGCTGCTGCTGCTGCTG 408
|||||
156 lyAspGluPheAsnAlaTyrTyrAlaArgArg 166
|||||
409 GTCGAGACCAAGGAGCAGATTCTGCCACCGC 440
```

seq\_name: /cgn2\_6/ptodata/1/lna/58\_COMB.seq:US-08-997-362-115

## seq\_documentation\_block:

Sequence 115, Application US/08997362  
Patent No. 5985287  
GENERAL INFORMATION:  
APPLICANT: Tan, Paul  
APPLICANT: Hiyama, Jun  
APPLICANT: Visser, Elizabeth  
APPLICANT: Skinner, Margot  
APPLICANT: Scott, Linda  
APPLICANT: Prestidge, Ross  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR  
TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS  
NUMBER OF SEQUENCES: 194  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Offices of Ann W. Speckman

STREET: 2601 Elliott Avenue, Suite 4185  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98121

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/997,362  
FILING DATE:  
CLASSIFICATION:

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970  
FILING DATE: June 12, 1997  
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347  
FILING DATE: August 29, 1996  
ATTORNEY/AGENT INFORMATION:

NAME: Sleath, Janet  
REGISTRATION NUMBER: 37,007  
REFERENCE/DOCKET NUMBER: 11000.1002C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-269-0565  
TELEFAX: 206-269-0563  
TELEX:

## INFORMATION FOR SEQ ID NO: 115:

SEQUENCE CHARACTERISTICS:  
LENGTH: 647 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic RNA  
US-08-997-362-115

## alignment\_scores:

Quality: 107.00 Length: 162  
Ratio: 1.259 Gaps: 9  
Percent Similarity: 52.469 Percent Identity: 30.864

## alignment\_block:

US-09-508-832-10 x US-08-997-362-115

Align seg 1/1 to: US-08-997-362-115 from: 1 to: 647

```
17 ArgGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPr 33
|||||
8 AGACAATTGGGTATGACGAAGAGCGCGCGTGGCTCGAGCGGGGCTC 57
|||||
33 OthrSerLeuGlnThrGluProGlnGlyAsnPro GluGlyAsnHisGly 49
|||||
58 AACGCCCTCGCAGAGC .....CCGTAAAGGTGACGTTGGG 92
|||||
50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
|||||
93 CCCGAGCGCTCGCAGACCTGCTGCTGGAGAAAGTGGGGCG .....CC 135
|||||
66 oProAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83
|||||
136 CCCAGCATCACCA...ACGATGGTGTGTCCATCGCCA..... 169
|||||
83 etArgArgSerLeuLeuSerArgSerSerSerGlyTyrPheSerPhe 99
|||||
170 ...AGGATCGAGCTGGAGGACCGCTACGAGAGATCGCGCTGAGCTGG 217
|||||
100 AspThrAspArgSerProAlaPrometSerCysAspLysSerThrGlnTh 116
|||||
218 ...TCAAAGAGTGGCGAGAGACGACGACGTCGCGGGGCGGCGCACC 264
|||||
116 rProSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlas 133
|||||
```



```
265 ACCACCGCCAGCGTCTGCTCAGGCTCTGGTTCGGAGGCGCTGGCGAA 314
133 erMetArgGlnAla.GluProAlaAspMetArgProGluIleTrp..... 147
:: :::: ||| ::::||| |||::: |||
315 CGTCGAGCGCGCGCAACCGCT.....CGGCTCAAGCGTGGCATCG 358
148 .....leAlaGlnGluLeuArgAfGileG 156
359 AGAAGGCTGTCAGGCTGTCACCGAGTGGCTGCTGAAGTCGGCCCAAGGAG 408
156 lyAspGluPheAsnAlaTyrtAlaArgArg 166
|| :::: ||| ::::||| |||::: |||
409 GTCGAGACCAAGGACGAGATTCTGCCACGC 440
```

seq\_name: /cgn2\_6/ptodata/1/lna/6B\_COMB.seq:US-09-095-855-115

seq\_documentation\_block:

; Sequence 115, Application US/09095855

; Patent No. 6160093

; GENERAL INFORMATION:

; APPLICANT: Tan, Paul

; APPLICANT: Visser, Elizabeth

; APPLICANT: Skinner, Margot

; APPLICANT: Prestidge, Ross

; TITLE OF INVENTION: Compounds and Methods for

; TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections

; NUMBER OF SEQUENCES: 208

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Law Offices of Ann W. Speckman

; STREET: 2601 Elliott Avenue, Suite 4185

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98121

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/095,855

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/705,347

; FILING DATE: 29-AUG-1996

; APPLICATION NUMBER: 08/873,970

; FILING DATE: 12-JUN-1997

; APPLICATION NUMBER: 08/997,362

; FILING DATE: 23-DEC-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Sleath, Janet

; REGISTRATION NUMBER: 37,007

; REFERENCE/DOCKET NUMBER: 11000.1002c3

; TELEPHONE: 206-269-0565

; TELEFAX: 206-269-0563

; TELEX:

; INFORMATION FOR SEQ ID NO: 115:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 647 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: Genomic RNA

US-09-095-855-115

alignment\_scores:

Quality: 107.00

Ratio: 1.259

Percent Similarity: 52.469

Length: 162

Gaps: 9

Percent Identity: 30.864

alignment\_block:

US-09-508-832-10 x US-09-095-855-115

Align seg 1/1 to: US-09-095-855-115 from: 1 to: 647

```
17 ArgGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPr 33
|||||::: ::::||| |||::: |||
8 AGACAATTGGTATGACGAAGAGCCCGCGCTGGCTCGAGCGGCGCTC 57
|||||::: ::::||| |||::: |||
33 oThrSerLeuGlnThrGluProGlnGlyAsnPro.GluGlyAsnHisGly 49
||| ||||| |||::: |||::: |||
58 AACGCCCTCGCAGCG.....CCGTAAGGTGACGTTGGG 92
|||||::: ::::||| |||::: |||
50 GlyClyGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
|||||::: |||::: |||
93 CCCGAAGGTGCGCAACGTCGTGCTGGAGAAGAAGTGGGCG.....CC 135
|||||::: ::::||| |||::: |||
66 OPtAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83
||| ||||| :::: |||||
136 CCCACGATCACCA...ACGATGGTGTGTCATGCCA..... 169
|||||::: ::::||| |||::: |||
83 etArgArgSerSerLeuLeuSerArgSerSerGlyTyrrPheSerPhe 99
|||||::: |||::: |||
170 ..AGGAGATCGAGCTGGAGGACCCCTACGAGAAGATCGCGCTGAGCTGG 217
|||||::: ::::||| |||::: |||
100 AspThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnTh 116
::: |||||::: ::::||| |||::: |||
218 ...TCAAAGAGGTGCGCAACGACGACGACGTCGCGGCGGCGACGCC 264
||| |||||::: ::::||| |||::: |||
116 rProSerProProCysGlnAlaPheAsnHisTyrrLeuSerAlaMetAla 133
||| |||||::: ::::||| |||::: |||
265 ACCACCCCAACCGTCTCGCTCAGGCTCTGGTTCGCGAAGCGCTCGCAA 314
||| |||||::: ::::||| |||::: |||
133 erMetArgGlnAla.GluProAlaAspMetArgProGluIleTrp..... 147
::: ||| ::::||| |||::: |||
315 CGTCGAGCGCGCGCAACCGCT.....CGGCTCAAGCGTGGCATCG 358
|||||::: ::::||| |||::: |||
148 .....leAlaGlnGluLeuArgAfGileG 156
||| |||||::: ::::||| |||::: |||
359 AGAAGGCTGTCGAGGCTGTCACCGAGTGGCTGCTGAAGTCGGCCCAAGGAG 408
||| |||||::: ::::||| |||::: |||
156 lyAspGluPheAsnAlaTyrtAlaArgArg 166
||| |||||::: ::::||| |||::: |||
409 GTCGAGACCAAGGACGAGATTCTGCCACGC 440
```

seq\_name: /cgn2\_6/ptodata/1/lna/6B\_COMB.seq:US-09-324-542-115

seq\_documentation\_block:

; Sequence 115, Application US/09124542

; Patent No. 6328978

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Tan, Paul L.J.

; APPLICANT: Prestidge, Ross

; TITLE OF INVENTION: Methods and Compounds for the Treatment

; TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders

; FILE REFERENCE: 11000.1007c1

; CURRENT APPLICATION NUMBER: US/09/324,542

; CURRENT FILING DATE: 1999-06-02

; EARLIER APPLICATION NUMBER: US 08/997,080

; EARLIER FILING DATE: 1997-12-23

; NUMBER OF SEQ ID NOS: 194

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 115

; LENGTH: 647

; TYPE: DNA

; ORGANISM: Mycobacterium vaccae

US-09-324-542-115

alignment\_scores:

Quality: 107.00

Ratio: 1.259

Length: 162

Gaps: 9

Percent Similarity: 52.469 Percent Identity: 30.864

## alignment\_block:

US-09-508-832-10 x US-09-324-542-115 ..

Align seg 1/1 to: US-09-324-542-115 from: 1 to: 647

```
17 ArgGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaApr 33
|||||:|||||:|||||:|||||:|||||:
8 AGACAAATGGGTATGACGAAGAGCGCCCGTGGCTCGAGCGGGGCTC 57
|||||:|||||:|||||:|||||:|||||:
33 oThrSerLeuGlnThrGluProGlnGlyAsnPro.GluGlyAsnHisGly 49
||| |||||:|||||:|||||:|||||:
58 AACGCCCTCCAGACG.....CCGTAAAGGTGACGTTGGG 92
|||||:|||||:|||||:|||||:
50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
|||||:|||||:|||||:|||||:|||||:
93 CCCGAAGGGTCCGACGCTGCTGCTGGAGAAGTGGGCG.....CC 135
|||||:|||||:|||||:|||||:
66 oProAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83
||| |||||:|||||:|||||:|||||:
136 CCCAGATACCA...ACGATGGTGTCTCATCGCCA.....169
|||||:|||||:|||||:|||||:
83 eAlaArgSerSerLeuLeuSerAlaArgSerSerGlyTyTyPheSerPhe 99
|||||:|||||:|||||:|||||:|||||:
170 ..AGGACATCGAGCTGGAGGACCGCTACGAGNAGATCGCGCTGACCTGG 217
|||||:|||||:|||||:|||||:
100 AspThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnH 116
||| |||||:|||||:|||||:|||||:
218 ...TCGAAGAGGTGCGCAAGACGACGACGCTGCGGGGCGGACGCC 264
|||||:|||||:|||||:|||||:
116 rProSerProCysGlnAlaPheAsnHisTyzLeuSerAlaMetAlaS 133
||| |||||:|||||:|||||:|||||:
265 ACCACCGCCACCGTCTGCTCAGGCTCTGCTCGCAAGGCGCTGCCA 314
||| |||||:|||||:|||||:|||||:
133 erMetArgGlnAla.GluProAlaAspMetArgProGluIleTrp..... 147
||| |||||:|||||:|||||:|||||:
315 GGTGCGAGCGCGCGCAACCGCT.....CGGCTCAAGCGTGGCATCG 358
|||||:|||||:|||||:|||||:
148 .....IleAlaGlnGluLeuArgArgIleg 156
||| |||||:|||||:|||||:|||||:
359 AGAAGGCTGTGAGGCTGTACCCAGTCTGCTGTAAGTGGGCCAAGGAG 408
||| |||||:|||||:|||||:|||||:
156 lyAspGluPheAsnAlaTyTyAlaArgArg 166
||| |||||:|||||:|||||:|||||:
409 GTCGAGACCAAGGACGAGATTTCTGCCACCGC 440
```

seq\_name: /cgn2\_6/plodata/1/ina/5B\_COMB.seq:US-08-602-359A-24

## seq\_documentation\_block:

; Sequence 24, Application US/08602359A

; Patent No. 5942430

## GENERAL INFORMATION:

; APPLICANT: ROBERTSON, Daniel E.

; APPLICANT: MURPHY, Dennis

; APPLICANT: REID, John

; APPLICANT: MAFFIA, Anthony

; APPLICANT: LINK, Steven

; APPLICANT: SWANSON, Ronald V.

; APPLICANT: WARREN, Patrick V.

; TITLE OF INVENTION: ESTERASES

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FISH & RICHARDSON P.C.

; STREET: 4225 EXECUTIVE SQUARE, STE 1400

; CITY: LA JOLLA

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 INCH DISKETTE

; COMPUTER: IBM PS/2

```
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602.359A
; FILING DATE: February 16, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: HAILE, LISA A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/010001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5070
; TELEFAX: 619-678-5099
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 NUCLEOTIDES
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: GENOMIC DNA
; US-08-602-359A-24

alignment_scores:
Quality: 107.00 Length: 177
Ratio: 1.244 Gaps: 9
Percent Similarity: 48.588 Percent Identity: 27.684
```

## alignment\_block:

US-09-508-832-10 x US-08-602-359A-24/rev ..

Align seg 1/1 to reverse of: US-08-602-359A-24 from: 1 to: 1041

```
2 AlalysGlnProSerAspValSerSerGluCysAsp.....13
|||||:|||||:|||||:|||||:
757 GCACACCCACCCACCGCGGCGGTCTTGTACTCCAGCCACCGGGGT 708
|||||:|||||:|||||:|||||:
14 ....ArgGluGlyArgGlnLeuGlnProAlaGluArgProGlnLeuA 29
||| |||||:|||||:|||||:|||||:
707 ATGACGTCCTCCGAGCCCTATAGTACGGGCTATCAGCAGCCCGCT...A 661
|||||:|||||:|||||:|||||:
29 rArgProGlyAla.....ProThrSerLeuGlnThrGluProGln 41
|||||:|||||:|||||:|||||:
660 GACCGCGGGTCCCGACCCCTCCACGATCGCTACAGCGCGGCCATAC 611
|||||:|||||:|||||:|||||:
42 GlyAsnProGluGlyAsnHisGlyGlyAspSerCysProHisG1 58
|||||:|||||:|||||:|||||:
610 TGAACCCCAATATTATCCGGCGCGGGGAAACCGCTCCGAGACATAG 561
|||||:|||||:|||||:|||||:
58 ySerProGlnGlyProLeuAlaProAlaSerProGlyProPheAla 75
||| |||||:|||||:|||||:|||||:
560 CCCACACAGGCC.....GGCATCCAGCACCTCCCGGGGCCAATCGT 517
|||||:|||||:|||||:|||||:
75 hrArgSerProLeuPheIlePheMetArgSerSerLeuLeuSerArg 91
||| |||||:|||||:|||||:|||||:
516 COTCGAGCCCC.....CGC 503
|||||:|||||:|||||:|||||:
92 SerSerSerGlyTyTyPheSer.....PheAspThrAs 102
||| |||||:|||||:|||||:|||||:
502 TCTCCCGTGGCCCGGAAGTCGAACACACCGGGTACCCCGCACATCC 453
|||||:|||||:|||||:|||||:
102 pArgSerProAlaProMetSerCysAspLysSerThrGln.ThrProSer 118
|||||:|||||:|||||:|||||:|||||:
452 ACGAGCTCCCGGGCCAGCACGACCATGTAGGGCGCGGAGCGGCCAGT 403
|||||:|||||:|||||:|||||:
119 ProProCysGlnAlaPheAsnHisTyzLeuSerAlaMetAlaSerMet 135
||| |||||:|||||:|||||:|||||:
402 ATACCCGTGCATCAAAACGAACACCGGGTTCGCCCGCGCGGGGCCA 353
|||||:|||||:|||||:|||||:
135 gGlnAlaGluPro.....AlaAspMetArg.....ProGluIle 147
```



```
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1569 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-997-362-113

alignment_scores:
  Quality: 107.00      Length: 162
  Ratio: 1.259        Gaps: 9
Percent Similarity: 52.469 Percent Identity: 30.864

alignment_block:
US-09-508-832-10 x US-08-997-362-113 ..

Align seg 1/1 to: US-08-997-362-113 from: 1 to: 1569

17 ArgGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPr 33
|||||
8 AGACAAATGGCGTATGACGAAGAGCCCGCGTGGCTCGAGCGGGGCTC 57
33 oThrSerLeuGlnThrGluProGlnGlyAsnPro.GluGlyAsnHisGly 49
|||
58 AACGCCCTCGCAGACG.....CCGTAAAGGTGACGTTGGG 92
50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
|||||
93 CCCGAAGGCGCGAAGCGTCTGCTGCGAAGAAGTGGCGCG.....CC 135
66 oProAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83
|||
136 CCCACGATCACCA...ACGATGGTGTCTCCATCGCCA.....169
83 etArgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPhe 99
|||||
170 ..AGAGATCGAGCTGGAGGACCGGTACGAGAGATCGCGCTGAGCTGG 217
100 AspThrAspArgSerProAlaProMetSerCysAspHisSerThrGlnTh 116
...
218 ...TCAAAGAGGTGCGCAAGACACGACGCTCGCGCGCGACGCCACC 264
116 rProSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlas 133
|||
265 ACCACCGCCACCGTCTGCTCAGGCTCTGCTCCGGAAGCGCTGCCAA 314
133 erMetArgGlnAla.GluProAlaAspMetArgProGluIleTrp..... 147
...
315 GTCGACAGCGCGCCCAACCCGCT.....CGGCTCAAGCGTGGCATCG 358
148 .....IleAlaGlnLeuLeuArgArgIleG 156
359 AGAAGGCTGTGAGGCTGTCAACCGTCTGCTGCTGAAGTCGGCAAGGAG 408
156 lyAspGluPheAsnAlaTyrTyrAlaArgArg 166
409 GTCGAGCAAGGAGGACGATTTCTGCCACCGC 440

seq_name: /cgn2_6/ptodata/1/1na/68_COMB.seq:US-09-095-855-113

seq_documentation_block:
; Sequence 113, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
```

```
; TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,855
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; APPLICATION NUMBER: 08/873,970
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: 08/997,362
; FILING DATE: 23-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1569 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-09-095-855-113

alignment_scores:
  Quality: 107.00      Length: 162
  Ratio: 1.259        Gaps: 9
Percent Similarity: 52.469 Percent Identity: 30.864

alignment_block:
US-09-508-832-10 x US-09-095-855-113 ..

Align seg 1/1 to: US-09-095-855-113 from: 1 to: 1569

17 ArgGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPr 33
|||||
8 AGACAAATGGCGTATGACGAAGAGCCCGCGTGGCTCGAGCGGGGCTC 57
33 oThrSerLeuGlnThrGluProGlnGlyAsnPro.GluGlyAsnHisGly 49
|||
58 AACGCCCTCGCAGACG.....CCGTAAAGGTGACGTTGGG 92
50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
|||||
93 CCCGAAGGCGCGAAGCGTCTGCTGCGAAGAAGTGGCGCG.....CC 135
66 oProAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83
|||
136 CCCACGATCACCA...ACGATGGTGTCTCCATCGCCA.....169
83 etArgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPhe 99
|||||
170 ..AGAGATCGAGCTGGAGGACCGGTACGAGAGATCGCGCTGAGCTGG 217
```

```

100 ASDThraspATGSerProAlaPrometSerCysAspLysSerThrGlnTh 116
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
218 ...TCAAGAGCTGCCCAAGAACCGACGCTCCGGCGGCGGCGGCGCACC 264
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
116 rProSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlas 133
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
265 ACCACCGCCACCGTGTCTGAGGCTCTGCTTCCGGAAGGCGCTGCGCAA 314
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
133 erMetArgGlnAla.GluProAlaAspMetArgProGluIleTrp.... 147
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
315 CGTCGACGCGCGGCCCAACCGCT.....CGGCCTCAAGCGTGGCATCG 358
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
148 .....IleAlaGlnGluLeuArgArgileG 156
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
359 AGAAGCGCTGTCGAGGCTGTCCACCGACGCTGCTGAAGTCGCGCAAGGAG 408
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
156 lyAspGluPheAsnAlaTyrTyrAlaArg 166
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
409 GTCGAGACCAAGGACAGATTCTGCCACCGC 440
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
seq_name: /cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-324-542-113

```

```

seq_documentation_block:
; Sequence 113, Application US/09324542
; Patent No. 5328978
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; FILE REFERENCE: 11000.1007c1
; CURRENT APPLICATION NUMBER: US/09/324,542
; CURRENT FILING DATE: 1999-06-02
; EARLIER APPLICATION NUMBER: US 08/997,080
; EARLIER FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; TYPE: DNA
; ORGANISM: Mycobacterium vaccae
US-09-324-542-113

```

```

alignment_scores:
  Quality: 107.00      Length: 162
  Ratio: 1.259        Gaps: 9
  Percent Similarity: 52.469  Percent Identity: 30.864

```

## alignment\_block:

US-09-508-832-10 x US-09-324-542-113 ..

Align seq 1/1 to: US-09-324-542-113 from: 1 to: 1569

```

17 ArgGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaPr 33
   |||||.....: : : : : : : : : : : : : : : : : :
8  ACACATTGCTATGACGAGAGCGCCGCGGCTGCGCTGCGCGCGGCGCTC 57
   |||||.....: : : : : : : : : : : : : : : : : :
33 othrSerLeuGlnThrGluProGlnGlyAsnPro.GluGlyAsnHisGly 49
   |||||.....: : : : : : : : : : : : : : : : : :
58 AACGCCCTCGGAGCG.....CCGTAAAGGTGACGTTGGG 92
   |||||.....: : : : : : : : : : : : : : : : : :
50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
   |||||.....: : : : : : : : : : : : : : : : : :
93 CCGGAAGGTCGACACGTCGCTGGGAGAGAGTGGGGCG.....CC 135
   |||||.....: : : : : : : : : : : : : : : : : :
66 oProAlaSerProGlyProPheAlaThrArgSerProPhePheIlePheM 83
   |||||.....: : : : : : : : : : : : : : : : : :
136 CCCACGATCACCA...ACGATGGTGTTCATCGCCA..... 169
83 eArgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPhe 99
   |||||.....: : : : : : : : : : : : : : : : : :

```

```

170 ..AGGACATCGAGCTGGAGGACCGCTACGAGAGATCGGCGCTGAGCTGG 217
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
100 AspThrAspArgSerProAlaPrometSerCysAspLysSerThrGlnTh 116
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
218 ...TCAAGAGCTGCCCAAGAACCGACGCTCCGGCGGCGGCGGCGCACC 264
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
116 rProSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlas 133
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
265 ACCACCGCCACCGTGTCTGAGGCTCTGCTTCCGGAAGGCGCTGCGCAA 314
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
133 erMetArgGlnAla.GluProAlaAspMetArgProGluIleTrp.... 147
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
315 CGTCGACGCGCGGCCCAACCGCT.....CGGCCTCAAGCGTGGCATCG 358
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
148 .....IleAlaGlnGluLeuArgArgileG 156
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
359 AGAAGCGCTGTCGAGGCTGTCCACCGACGCTGCTGAAGTCGCGCAAGGAG 408
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
156 lyAspGluPheAsnAlaTyrTyrAlaArg 166
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
409 GTCGAGACCAAGGACAGATTCTGCCACCGC 440
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

seq\_name: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:US-08-997-080-159

## seq\_documentation\_block:

```

; Sequence 159, Application US/08997080
; Patent No. 5968524
; GENERAL INFORMATION:
; APPLICANT: WATSON, JAMES D.
; APPLICANT: TAN, PAUL L.J.
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,080
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 159:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1626 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-997-080-159

```

```

alignment_scores:
  Quality: 107.00      Length: 162
  Ratio: 1.259        Gaps: 9
  Percent Similarity: 52.469  Percent Identity: 30.864

```

## alignment\_block:

US-09-508-832-10 x US-08-997-080-159

Align seg 1/1 to: US-08-997-080-159 from: 1 to: 1626

```
17 ArgGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPr 33
|||||
8 AGACAAATTCGTATGACGAAGAGCCCGCTGCGCTCGAGCGGGCCTC 57
|||||
33 oThrSerLeuGlnThrGluProGlnGlyAsnPro.GluGlyAsnHisGly 49
|||
58 AACGCCCTCGCAGACG.....CCGTAAAGGTGACGTTGGG 92
|||||
50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
|||||
93 CCCGAAGGGTCGACGCTGCTCGAGAGAAAGTGGGCG.....CC 135
|||||
66 oProAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83
|||||
136 CCCACGATCACCA...ACGATGCTGTCTCCATCCCA..... 169
|||||
83 eTArgArgSerLeuLeuSerArgSerSerGlyTyrPheSerPhe 99
|||||
170 ..AGGAGATCGACGTCGAGGACCCGTACGAGAACATCGCGCTGAGCTGG 217
|||||
100 AspThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnTh 116
|||||
218 ...TCAAAGAGTCCGACGAGACCGACGCTCGCGGCGAGCGCAC 264
|||||
116 rProSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlas 133
|||||
265 ACCACGCCACCGCTGCTCGAGCTCTGCTCGGAGGCGCTGGCGCA 314
|||||
133 erMetArgGlnAla.GluProAlaAspMetArgProGluIleTyrPhe 147
|||||
315 CGTCGACGCGCGCAACCCGCT.....CGGCTCAACGCTGGCATCG 358
|||||
148 .....IleAlaGlnGluLeuArgArgIleg 156
|||||
359 AGAGGGCTGTCGAGGCTGTACCCAGTCGCTGTAAGTCGCGCAAGGAG 408
|||||
156 lyAspGluPheAsnAlaTyrTyrAlaArgArg 166
|||
409 GTCGAGACCAAGAGCAGATTTCTGCCACCGC 440
|||||
```

seq\_name: /cgn2\_6/ptodata/1/ins/5B\_COMB.seq:us-08-997-362-159

## seq\_documentation\_block:

Sequence 159: Application US/08997362

Patent No. 5985287

## GENERAL INFORMATION:

APPLICANT: Tan, Paul  
APPLICANT: Hiyama, Jun  
APPLICANT: Visser, Elizabeth  
APPLICANT: Skinner, Margot  
APPLICANT: Scott, Linda  
APPLICANT: Prestidge, Ross

## TITLE OF INVENTION: COMPOUNDS AND METHODS FOR

## NUMBER OF SEQUENCES: 194

## CORRESPONDENCE ADDRESS:

ADDRESSES: Law Offices of Ann W. Speckman

STREET: 2601 Elliott Avenue, Suite 4185

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98121

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASCTEQ for Windows Version 2.0

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/997.362  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970  
FILING DATE: June 12, 1997  
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347  
FILING DATE: August 29, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Sleath, Janet  
REGISTRATION NUMBER: 37,007  
REFERENCE/DOCKET NUMBER: 11000.1002C2  
TELEPHONE: 206-269-0565  
TELEFAX: 206-269-0563  
TELEX:  
INFORMATION FOR SEQ ID NO: 159:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1626 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-997-362-159

## alignment\_scores:

Quality: 107.00 Length: 162

Ratio: 1.259 Gaps: 9

Percent Similarity: 52.469 Percent Identity: 30.864

## alignment\_block:

US-09-508-832-10 x US-08-997-362-159

Align seg 1/1 to: US-08-997-362-159 from: 1 to: 1626

```
17 ArgGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPr 33
|||||
8 AGACAAATTCGTATGACGAAGAGCCCGCTGCGCTCGAGCGGGCCTC 57
|||||
33 oThrSerLeuGlnThrGluProGlnGlyAsnPro.GluGlyAsnHisGly 49
|||
58 AACGCCCTCGCAGACG.....CCGTAAAGGTGACGTTGGG 92
|||||
50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
|||||
93 CCCGAAGGGTCGACGCTGCTCGAGAGAAAGTGGGCG.....CC 135
|||||
66 oProAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83
|||||
136 CCCACGATCACCA...ACGATGCTGTCTCCATCCCA..... 169
|||||
83 eTArgArgSerLeuLeuSerArgSerSerGlyTyrPheSerPhe 99
|||||
170 ..AGGAGATCGACGTCGAGGACCCGTACGAGAACATCGCGCTGAGCTGG 217
|||||
100 AspThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnTh 116
|||||
218 ...TCAAAGAGTCCGACGAGACCGACGCTCGCGGCGAGCGCAC 264
|||||
116 rProSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlas 133
|||||
265 ACCACGCCACCGCTGCTCGAGCTCTGCTCGGAGGCGCTGGCGCA 314
|||||
133 erMetArgGlnAla.GluProAlaAspMetArgProGluIleTyrPhe 147
|||||
315 CGTCGACGCGCGCAACCCGCT.....CGGCTCAACGCTGGCATCG 358
|||||
148 .....IleAlaGlnGluLeuArgArgIleg 156
|||||
359 AGAGGGCTGTCGAGGCTGTACCCAGTCGCTGTAAGTCGCGCAAGGAG 408
|||||
156 lyAspGluPheAsnAlaTyrTyrAlaArgArg 166
|||
```

409 GTCGAGACCAAGGAGCAGATTTCGCCACCGC 440

seq\_name: /cgn2\_6/ptodata/1/lna/6B\_COMB.seq:US-09-095-855-159

seq\_documentation\_block:

Sequence 159, Application US/09095855

Patent No. 6160093

GENERAL INFORMATION:

APPLICANT: Tan, Paul

APPLICANT: Visser, Elizabeth

APPLICANT: Skinner, Margot

APPLICANT: Prestidge, Ross

TITLE OF INVENTION: Compounds and Methods for

TITLE OF INVENTION: Treatment and Diagnosis of Nycobacterial Infections

NUMBER OF SEQUENCES: 208

CORRESPONDENCE ADDRESS:

ADDRESSEE: Law Offices of Ann W. Speckman

STREET: 2601 Elliott Avenue, Suite 4185

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98121

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/095,855

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/705,347

FILING DATE: 29-AUG-1996

APPLICATION NUMBER: 08/873,970

FILING DATE: 12-JUN-1997

APPLICATION NUMBER: 08/997,362

FILING DATE: 23-DEC-1997

ATTORNEY/AGENT INFORMATION:

NAME: Sleath, Janet

REGISTRATION NUMBER: 37,007

REFERENCE/DOCKET NUMBER: 11000.1002c3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-269-0565

TELEFAX: 206-269-0563

TELEX:

INFORMATION FOR SEQ ID NO: 159:

SEQUENCE CHARACTERISTICS:

LENGTH: 1626 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-095-855-159

alignment\_scores:

Quality: 107.00 Length: 162

Ratio: 1.259 Gaps: 9

Percent Similarity: 52.469 Percent Identity: 30.864

alignment\_block:

US-09-508-832-10 x US-09-095-855-159 ..

Align seg 1/1 to: US-09-095-855-159 from: 1 to: 1626

17 ArgGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPr 33

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

8 AGACAATTGCGTATGACGAGAGAGCGCGCGTGGCCCTCGAGCGGGCCCTC 57

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

33 oThrSerLeuGlnThrGluProGlnGlyAsnPro.GluGlyAsnHisGly 49

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

58 AACGCCCTCGCAGACG.....CCGTAAAGGTGACGCTTGGG 92

50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
93 CCGAAGAGGTGCGAACCGTGTGGAGAGAACTGGGGCG.....CC 135  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
66 oProAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
136 CCACGATCACCA...ACGATGGTGTGTCTCCATCGCCA..... 169  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
83 etArgArgSerSerLeuLeuSerArgSerSerGlyTyPheSerPhe 99  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
170 ..AGGAGATCGAGCTGGAGGACCGCGTACGAGAGATCGGCGCTGAGCTGG 217  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
100 AspThrAspArgSerProAlaProMetSerCysAspSlySerThrGlnTh 116  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
218 ...TCAAAGAGGTGCGCCAGAGACCGACGCTGCGCGCGCGACGCCACC 264  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
116 rProSerProCysGlnAlaPheAsnHisTyLeuSerAlaMetAla 133  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
265 ACCACCGCCACCGCTGCTCGCTCAGGCTCTGCTCCGAGGCGCTGGCGAA 314  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
133 erMetArgGlnAla.GluProAlaAspMetArgProGluIleTrp..... 147  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
315 CGTCGACGCGCGCGCCACACCGCT.....CGGCTCAAGCGTGGCATCG 358  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
148 .....IleAlaGlnGluLeuArgArgIleG 156  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
359 AGAAGGCTGTGCGAGGCTGTACCCCGAGTGTGCTGTAAGTCGCCCAAGGAG 408  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
156 lYAspGluPheAsnAlaTyTyAlaArgArg 166  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
409 GTCGAGACCAAGGAGCAGAGATTTCGCCACCGC 440









## ORIGIN

alignment\_scores:      Quality: 667.00      Length: 146  
                             Ratio: 4.869      Gaps: 2  
                             Percent Similarity: 93.836      Percent Identity: 89.726

alignment\_block:  
 US-09-508-832-10 x BB651225      ..

Align seg 1/1 to: BB651225 from: 1 to: 645

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGly 17  
 |||||  
 220 ATGCCAAGCAACCTTCGTATGTAAGTCTGAGTGTGACAGAAAGGTGG 269  
 |||||  
 17 -GlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34  
 |||||  
 270 ACAATTGCAGCCTGCTGACAGAGCCCTCCCGAGCTCAGGCGCTGGGCCCTTA 319  
 |||||  
 34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50  
 |||||  
 320 CCTCCCTACAGACAGAACCGCAGGTAATCCGAC.....GGC 357  
 |||||  
 51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67  
 |||||  
 358 GAAGGGAGCGCTGCCCCCGCCAGCGCAGCCCTCAGGGGCCCGCTGCCGCCACC 407  
 |||||  
 67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMetA 84  
 |||||  
 408 GGCCAGCCCTGGCCCTTTTGTCTACAGATCCCGACATTTTCATCTTTGTGA 457  
 |||||  
 84 rGArgSerSerLeuSerArgSerSerGlyTyrPheSerPheAsp 100  
 |||||  
 458 GAAGRTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 507  
 |||||  
 101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117  
 |||||  
 508 ACAGACAGAGCGCCGCGCCAGCCAGTGTGACAAAGTCAACACAAACCCC 557  
 |||||  
 117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134  
 |||||  
 558 AAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 606  
 |||||  
 134 eTArgGlnAla.....GluProAlaAspMetArgPro 144  
 |||||  
 607 TACGACAGTCTCAGGAGGAACCTGAAGATCTGCGCCCG 644

seq\_name: gb\_gss:A2706148

seq\_documentation\_block:

LOCUS      A2706148      580 bp      DNA      linear      GSS 24-JAN-2001

DEFINITION      RPCI-23-227P3-TV RPCI-23 Mus musculus genomic clone RPCI-23-227P3,

DNA sequence.

ACCESSION      A2706148

VERSION      A2706148.1      GI:12433319

KEYWORDS      GSS.

SOURCE      house mouse.

## ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 580)

AUTHORS      Zhao, S., Niemman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret

B., Levins, M., McGann, S., Tsagaye, G., Geer, K., Krol, M., de Jong, P.

and Fraser, C.M.

Mouse BAC End Sequences from Library RPCI-23

Unpublished (1999)

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200  
 Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end page: [http://www.tigr.org/tldb/bac\\_ends/mouse/bac\\_ends\\_intro.html](http://www.tigr.org/tldb/bac_ends/mouse/bac_ends_intro.html)

Plate: 227      row: P      column: 3

Seq primer: 17

Class: BAC ends.

Location/Qualifiers

FEATURES  
 source

1..580  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCI-23-227P3"  
 /clone\_lib="RPCI-23"  
 /sex="Female"  
 /lab\_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site: 1;  
 ECORI: Site 2; ECORI: Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of ECORI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at the ECORI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT      138 a      162 c      138 g      142 t

ORIGIN

alignment\_scores:

Quality: 656.00      Length: 131  
 Ratio: 5.248      Gaps: 1  
 Percent Similarity: 95.420      Percent Identity: 93.893

alignment\_block:

US-09-508-832-10 x A2706148      ..

Align seg 1/1 to: A2706148 from: 1 to: 580

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGly 17

90 ATGCCAAGCAACCTTCGTATGTAAGTCTGAGTGTGACAGAAAGGTGG 139

17 gClnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34

140 ACAATTGCAGCCTGCTGACAGAGCCCTCCCGAGCTCAGGCGCTGGGCCCTTA 189

34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50

190 CCTCCCTACAGACAGAACCGCAGGTAATCCGAC.....GGC 227

51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67

228 GAAGGGAGCGCTGCCCCCGCCAGCGCCTCAGGGCCCGCTGGGCCACC 277

67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMetA 84

278 GGCCAGCCCTGGCCCTTTTGTCTACAGATCCCGACATTTTCATCTTTGTGA 327

84 rGArgSerSerLeuSerArgSerSerGlyTyrPheSerPheAsp 100

328 GAAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 377

101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117

378 ACAGACAGAGCGCCGCGCCAGTGTGACAAAGTCAACACAAACCCC 427

117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMet 131

428 AAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 470

```

seq_name: gb_est1:BB577135
seq_documentation_block:
LOCUS      BB577135          586 bp      mRNA      linear      EST 26-OCT-2001
DEFINITION Musculus riken full-length enriched, 13 days embryo male testis Mus
ACCESSION  BB577135
VERSION    BB577135.2 GI:16449433
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 686)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Hori,Y., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakai,H., Sasaki
D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
UNPUBLISHED (2001)
On Nov 30, 2000 this sequence version replaced gi:11473679.
Contact: Yoshinhide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,M., Sugahara,Y., Shibata,K., Itoh
M., Komoto,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Komoto,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
FEATURES             Location/Qualifiers
     source          1..686
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone_lib="RIKEN full-length enriched, 13 days embryo
                     male testis"
                     /sex="male"
                     /tissue_type="testis"
                     /dev_stage="13 days embryo"
/lab_host="DH10B"
/note="Site_1: Sall; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGACGACGCTCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 5.0 and subtraction to Rot = 100.0. Second strand
cDNA was prepared with the primer adapter of sequence [5'
GAGAGAGACGATTCGAGTTAATTAATCCGCCGCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: Sall; 3' end: BamHI."
BASE COUNT      145 a 228 c 167 g 146 t
ORIGIN
alignment_scores:
    Quality: 653.00      Length: 131
    Ratio: 5.224        Gaps: 1
    Percent Similarity: 95.420 Percent Identity: 93.130
alignment_block:
US-09-508-832-10 x BB577135 ..
Align seq 1/1 to: BB577135 from: 1 to: 686
1 MetAlaLySGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
220 ATGGCCAAACCAACCTTCGATGATAGTTCAGTGCAGACAGAAAGGTGG 269
17 gGlnLeuGlnProAlaGluArGProProGlnLeuArGProGlyAlaProT 34
270 ACAATTCCAGCCCTGCTGAGAGGGCCCTCCAGCTCAGGCCCTGGGGCCCTA 319
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
320 CCTCCCTACAGACAGACGCCGCAAGCTAATCCGAC.....GGC 357
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
358 GAAGGGGACCGCTGGCCGCCAGCGCCCTCAGGGCCCTGCGCCGCCACC 407
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeta 84
408 GGCCAGCCCTGGCCCTTTTGTCTACAGATCCCACTTTTCATCTTTGTGA 457
84 rArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 100
458 GAAGATCTCTCTGCTGCTCCCGCTCCAGTGGGTATTTCTCTTTGAC 507
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
508 ACAGACAGAGGCCCGGCCACCCATGAATTTGTGACAAAGTCAACACAAACCCC 557
117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMet 131
558 AAGTCCTCTCTTGGACGCCCTTCAACCACTATCTCATGTGCAATG 600
seq_name: gb_est2:BI906766
seq_documentation_block:
LOCUS      BI906766          700 bp      mRNA      linear      EST 16-OCT-2001
DEFINITION 603064506F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5213713 5',
KEYWORDS   mRNA sequence.
ACCESSION  BI906766
VERSION    BI906766.1 GI:16169524
KEYWORDS   EST.

```

SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 700)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L14M1536 row: 1 column: 02  
High quality sequence stop: 696.  
Location/Qualifiers  
1. .700  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5213713"  
/clone\_lib="NIH\_MGC\_118"  
/tissue\_type="leukocyte"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV  
(destroyed); RNA source leukocytes from anonymous pool of  
non-activated adult donors. Library is oligo-dT primed  
and directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.7 kb, insert size range  
1.2-3.3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 027. Note:  
this is a NIH MGC library."

BASE COUNT 148 a 233 c 175 g 144 t  
ORIGIN

alignment\_scores:  
Quality: 651.50 Length: 151  
Ratio: 4.402 Gaps: 2  
Percent Similarity: 98.013 Percent Identity: 95.364

alignment\_block:  
US-09-508-832-10 x B1906766 ..

Align seg 1/1 to: B1906766 from: 1 to: 700

1 MetAlaLysGlnProSerAspValSerSer.GluCysAspArgGluGlyA 17  
|||||  
257 ATGGCAAGCAAGCCTTCTGATGTAAGTTCTGAGTGTGACCGAGAGGTA 306  
|||||  
17 rGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPro 33  
|||||  
307 GACAAATTGCAGCGTCGCGAGAGCGCTCCCGAGCTCAGACCTGGGGCCCT 356  
|||||  
34 ThrSerLeuGlnThrGlnProGlnGlyAsnProGluGlyAsnHisGlyGI 50  
|||||  
357 ACCCTCCCTACAGACAGACGACCAAGGTAATCTGAAGCAATACAGGAGG 406  
|||||  
50 yClnGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProp 67  
|||||  
407 TCAAGGGGACAGCTGCCCGCCAGGAGCC...TCAGGGCCCGTGCCCGCAC 453  
|||||  
67 roAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMet 83  
|||||  
454 CTGCGACCCCTGG. CTTTGTGCTACAGATCCCGCTTTTCATCTTATG 502  
|||||  
84 ArgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheAs 100  
|||||  
503 AGAAGATCCCTCCCTGCTGCTGATCTCCAGTGGGTATTC.TCTTTTGA 551  
|||||

100 pThrAspArgSerProAlaPrometSerCysAspLysSerThrGlnThrP 117  
|||||  
552 CACACAGACGAGCCACACCCATGAGTGTGACAAATCACACACACCC 601  
|||||  
117 roSerProProCysGluAlaPheAsnHisTyrLeuSerAlaMetAlaSer 133  
|||||  
602 CAAGTCCTCTTGGCAGGCTTCAA.CACTATCTCAGTCAATGGCTTCC 650  
|||||  
134 MetArgGlnAlaGluProAlaAspMetArgProGluIleTyrPheAlaG 150  
|||||  
651 ATGAGGACGGCTGAACCTGGAGATATGCG.CCAGAGATATGATCGCCCA 699  
|||||  
150 n 150  
700 G 700

seq\_name: gb\_est2:BF021882

seq\_documentation\_block:  
LOCUS BF021882 452 bp mRNA linear EST 29-DEC-2000  
DEFINITION uy59b09.y1 McCartney Eddy round spermatid Mus musculus cDNA clone  
IMAGE:3663833 5' similar to TR:054918 054918 BCL2 INTERACTING  
MEDIAN OF CELL DEATH ;, mRNA sequence.  
ACCESSION BF021882  
VERSION BF021882.1 GI:10753214  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 452)  
AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,  
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, I., Person,  
B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Rifter,  
E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,  
Waterston, R., and Wilson, R.  
The WashU-NCI Mouse EST Project 1999  
Contact: Marra M/WashU-NCI Mouse EST Project 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: [mouseestat@wustl.edu](mailto:mouseestat@wustl.edu)  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
MGI:1424601  
Seq primer: Primer name ambiguous  
High quality sequence stop: 386.  
Location/Qualifiers  
1. .452  
/organism="Mus musculus"  
/strain="CD-1"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3663833"  
/clone\_lib="McCarrey Eddy round spermatid"  
/sex="male"  
/tissue\_type="round spermatids, pooled from multiple mice"  
/dev\_stage="60 day"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: testis; Vector: pBluescript SK+ (Stratagene  
); Site\_1: XhoI; Site\_2: EcoRI; cDNA oligo dt-primed  
15'-(CA)10-ACTAGTCTCAGTCTTTT-3' and directionally  
cloned using 5' linkers 5'-AATTGGCAGCAG-3' and  
5'-CTCGTGGC-3'. Size selection of >400bp material gives  
average insert size ranging from 1-2 kb. Library was mass  
excised (from lambda-unizap-XR) and resulting  
single-stranded phagemids were prepped and transformed  
into DH10B. Library contains 98.5% recombinants.  
References: J. Androl. 20:635-639 and Gene 25:263-269.  
Library constructed and donated by J. McCarrey, Ph.D.  
(Southwest Foundation for Biomedical Research, Dept. of

Genetics): excision done by E.M. Eddy, Ph.D. (National Institutes of Health, National Institute of Environmental Health Sciences). Original lambda-based library is available through ATCC, catalog #63423.

BASE COUNT 106 a 130 c 112 g 104 t

ORIGIN

alignment\_scores:  
Quality: 622.00 Length: 139  
Ratio: 4.822 Gaps: 1  
Percent Similarity: 92.806 Percent Identity: 85.612

alignment\_block:

US-09-508-832-10 x BF021882 ..

Align seg 1/1 to: BF021882 from: 1 to: 452

```

62 GlyProLeuAlaProProAlaSerProGlyProPheAlaThrArgSerPr 78
|||||
1 GGCCCGCTGGCCGCCACCGCGCAGCCCTGGCCCTTTGCTACGAGATCCCC 50
|||||
78 oLeuPheMetArgSerSerLeuLeuSerArgSerSerG 95
|||||
51 ACTTTTCATCTTGTGAGAGATCTTCTGCTGCTCCCGGCTCTCCAGTG 100
|||||
95 lyTyrPheSerPheAspThrAspArgSerProAlaProMetSerCysAsp 111
|||||
101 GGTATTTCTCTTTTGACACAGAGAGCGCGGCACCATGAGTGTGAC 150
|||||
112 LysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTyrLe 128
|||||
151 AAGTCAACACAAACCCCAAGTCTCTCTGCCAGGCTTCAACCACTATCT 200
|||||
128 uSerAlaMetAlaSerMetArgGlnala.....GluProAlaAspMetA 143
|||||
201 CAGTGCATGCTTCCATACGACGCTCTCAGGAGAACCTGAGATCTGC 250
|||||
143 rProGluLeuTrpPheAlaGlnGluLeuArgArgIleGlyAspGluPhe 159
|||||
251 GCCCGGAGATACGGATTCACAGAGAGCTCGCGCGGATCGGAGAGAGTTC 300
|||||
160 AsnAlaTyrTyrAlaArgArgValPheLeuAsnAsnTyrGlnAlaAlaG 176
|||||
301 AACGAACTTACACAGAGAGGCTTTTGCATGATTTACCGCGAGGCTGA 350
|||||
176 uAspHisProArgMetValIleLeuArgLeuLeuArgTyrIleValArgL 193
|||||
351 AGACCCACCTCAATGCTTATCTTACAACTGTACGCTTATCTTCCTGC 400
|||||
193 euValTrpArgMetHis 198
|||||
401 TGGTATGGAAGGCAT 417

```

seq\_name: gb\_est2:BI454637

seq\_documentation\_block:

LOCUS BI454637 765 bp mRNA linear EST 21-AUG-2001  
DEFINITION 603170532F1 NCI\_CCAP\_Mam5 Mus musculus cDNA clone IMAGE:5250271 5', mRNA sequence.

ACCESSION BI454637  
VERSION BI454637.1 GI:15245293

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 765)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabps-re@mail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: InCyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM11631 row: 0 column: 08  
High quality sequence stop: 760.

FEATURES  
source

Location/Qualifiers  
1..765  
/organism="Mus musculus"  
/strain="C57/B6"  
/db\_xref="taxon:10090"  
/clone\_lib="NCI\_CCAP\_Mam5"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt:  
site: 2; NotI: Cloned unidirectionally. Primer: Oligo dr.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Robin Humphreys,  
NIH"

BASE COUNT 175 a 214 c 227 g 149 t  
ORIGIN

alignment\_scores:  
Quality: 576.00 Length: 201  
Ratio: 4.465 Gaps: 3  
Percent Similarity: 64.179 Percent Identity: 59.701

alignment\_block:

US-09-508-832-10 x BI454637 ..

Align seg 1/1 to: BI454637 from: 1 to: 765

```

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
|||||
172 ATGCCCAAGCAACCTTCTGATGTAAGTCTGAGTGTGACAGAGAGTGG 221
|||||
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyValaProT 34
|||||
222 ACANTTGCAGCCCTGCTGAGAGCCCTCCCACTCAGCCCTGGGGCCCCPA 271
|||||
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
|||||
272 CCTCCCTACAGACAGAACGCCAA..... 294
|||||
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
294 ..... 294
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMetA 84
294 ..... 294
84 rgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheAsp 100
294 ..... 294
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
|||||
295 ...GACAGAGCCCGCACCCATGATGTTGTGCAAGTCAACAAACCCC 341
|||||
117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134
|||||
342 AAGTCTCTCTGCCAGGCTTCAACCACTATCTCAGTGCATGCTTCCA 391
|||||
134 etArgGlnAla.....GluProAlaAspMetArgProGluIleTrpIle 148
|||||
392 TAGCAGACTCTCAGGAGGAACCTGAAGATCTCGCGCGGAGATACCGATT 441

```

149 AlaGlnLeuLeuArgIleGly AspGluPheAsnAlaTyrTyrAlaa 165  
 |||||  
 442 GCACAGAGCTCGGGAGTGGAGCGAGGTTCAACGAAACTTACACAA 491  
 |||||  
 165 rGargvalPheLeuAsnAsnTyrGlnAlaGluAspHisProArgMet 181  
 |||||  
 492 GGAGGGTGTTCGAATGATTACCGGAGGCTGAAGACCACCTCAAAATG 541  
 |||||  
 182 ValIleLeuArgLeuLeuArgTyrIleValArgLeuValIlePargMetH1 198  
 |||||  
 542 GTTATCTTACAACTGTACGCTTATCTTCGCTGATGGAGAGGCA 591  
 |||||

seq\_name: gb\_est2:BG173095

seq\_documentation\_block: 668 bp mRNA linear EST 06-FEB-2001  
 LOCUS BG173095 602336666F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:4459720 5',  
 DEFINITION mRNA sequence.

ACCESSION BG173095.1 GI:12679707

VERSION BG173095.1

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 668) Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [rcs@nih.nih.gov](mailto:rcs@nih.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

plate: LLAM10260 row: c column: 17

High quality sequence stop: 599.

Location/Qualifiers

1..668

/organism="Mus musculus"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:4459720"

/clone\_lib="NCI\_CGAP\_Mam1"

/tissue\_type="tumor, biopsy sample"

/dev\_stage="10 months, virgin"

/lab\_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;

Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

BASE COUNT 135 a 235 c 162 g 136 t

ORIGIN

alignment\_scores:

Quality: 544.50 Length: 154

Ratio: 4.125 Gaps: 8

Percent Similarity: 85.714 Percent Identity: 77.273

alignment\_block:

US-09-508-832-10 x BG173095

Align seg 1/1 to: BG173095 from: 1 to: 668

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17  
 |||||  
 223 ATGCCCAAGCAACCTTCTGATGTAGTGTGAGTGTGACAGAGAGGTGG 272  
 |||||  
 17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34  
 |||||  
 273 ACAATTCACGCTCGCTCAGAGGCTCCCGCAGCTCAGGCTGGGGCCCTA 322  
 |||||  
 34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50  
 |||||  
 323 CCTCCCTACAGACAGAACCCCAAGGTAAATCCGAC .....GGC 360  
 |||||  
 51 GluGlyAsp\_SerCysProHisGlySerProGlnGlyProLeuAlaProp 67  
 |||||  
 361 GAAGGGACCTGTGCTGGCCCCCAGGCGCTCAGGGCCCTGCTGGCCCCA 410  
 |||||  
 67 roAlaSerProGlyProPheAlaThrArgSer\_ProLeuPheHellePheMe 83  
 :::::  
 411 CCGGACGCTCGGCCCTTTTGCTACCCAGATCCGCCACCTTTTCATCTTGT 460  
 |||||  
 83 tarGArgSerSerLeuLeuSerArgSerSerGlyTyrPheSer...P 99  
 :|||  
 461 CAGAAATATCTTCTCTCTCTCCGCTCTCCAGTGGGTATATCTCCTT 510  
 |||||  
 99 heAspThrAspArgSerProAlaProMetSerCysAspLysSerThrGln 115  
 |||||  
 511 TTGACACAGCACAGGAGCGCCGCCACCCATGAGTGTGACAAGTCAACCAA 560  
 |||||  
 116 ThrProSerProProCysGlnAlaPheAsnHisTyrLeuSer\_AlaMet 131  
 |||||  
 561 AACCCCAAGTCTCTCTGCGGAGGCTTCAACCACTATCTCAGTTGCAATG 610  
 |||||  
 132 AlaSerMetArgGln.....Ala\_GluProAlaAspMetArgProGluI 146  
 :|||  
 611 GCTTTCATACAGACAGTCTCCAGGAGGAACCTGAGGATCTGCGCCCGGAGA 660  
 |||||

seq\_name: gb\_est1:BB653788

seq\_documentation\_block:

LOCUS BB653788 584 bp mRNA linear EST 26-OCT-2001

DEFINITION BB653788 RIKEN full-length enriched, adult male liver tumor Mus

musculus cDNA clone C730026B05 5', mRNA sequence.

ACCESSION BB653788.1 GI:16487616

VERSION BB653788

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 584)

AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,

Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda

,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Sakai

,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,

Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,

Muramatsu,M. and Hayashizaki,Y.

RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

Unpublished (2001)

JOURNAL

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: [genome-res@gsc.riken.go.jp](mailto:genome-res@gsc.riken.go.jp)

URL: <http://genome.gsc.riken.go.jp/>

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh





Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-resgsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura  
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
 Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.  
 e mouse tissues.

FEATURES  
 source  
 1. 480  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="G37003SM16"  
 /clone\_lib="RIKEN full-length enriched, B16 F10Y cells"  
 /cell\_type="B16 F10Y cells"  
 88 a 175 c 121 g 96 t

BASE COUNT  
 ORIGIN

alignment\_scores  
 Quality: 376.00 Length: 88  
 Ratio: 4.700 Gaps: 2  
 Percent Similarity: 50.909 Percent Identity: 88.636

alignment\_block:  
 US-09-508-832-10 x BB856627 ..

Align seg 1/1 to: BB856627 from: 1 to: 480

1 MetAlatysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17  
 |||||  
 230 ATGCCAAGCAACCTTCTGATGTTAGTTCTGAGTGTGACAGAGAGGTGG 279  
 |||||  
 17 gGlnLeuGlnProAla.GluArgProProGlnLeuArgProGlyAlaPro 33  
 |||||  
 280 ACAATTGCAGCTGCTGTAGAGGCTCCCGAGCTCAGGCTGCGGCCCT 329  
 |||||  
 34 ThrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyG 50  
 |||||  
 330 ACCTCCCTACAGACAGACG.CAAGGTAAATCCGAC.....GG 366  
 |||||  
 50 yGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProp 67  
 |||||  
 367 CGAAGGGGACCGCTGGCCCCGAGGCGCCCTCAGGGCCGCTGCCCCC 416  
 |||||  
 67 roAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMet 83  
 |||||  
 417 CGGCCAGCCCTGGCCCTTTTGTGTACAGATCCCGACCTTTTCATCTT 466  
 |||||  
 84 ArgArgSerSer 87  
 |||||  
 467 AGAAGATCTTCT 478  
 |||||

seq\_name: gb\_est1:BB631162  
 seq\_documentation\_block:

LOCUS  
 DEFINITION  
 BB631162 RIKEN full-length enriched, 16 days neonate thymus Mus  
 musculus cDNA clone A130059E14 5', mRNA sequence.  
 BB631162  
 BB631162.1 GI:16467982  
 EST.  
 house mouse.  
 Mus musculus  
 Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 1 (bases 1 to 620)  
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,  
 Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda  
 M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okada  
 Okazaki, Y., Okido, T., Saito, K., Sakai, C., Sakai, K., Sano, H., Sasaki  
 T., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,  
 Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,  
 Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)  
 Unpublished (2001)  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-resgsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura  
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
 Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Alzawa  
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
 Hayashizaki, Y.  
 Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.  
 e mouse tissues.

FEATURES  
 source  
 1. 620  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="A130059E14"  
 /clone\_lib="RIKEN full-length enriched, 16 days neonate  
 thymus"  
 /tissue\_type="thymus"  
 /dev\_stage="16 days neonate"  
 /lab\_host="DH10B"  
 /note="Site 1: Sali; Site 2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'  
 GAGAGAGAGAGGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse



117 oserProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134  
 |||||  
 431 AAGTCTCTTCCAGCCCTTCAACCACTATCTCAGTGCATGCATCAGT 480  
 |||||  
 134 eTArgGlnAlaGluProAlaAspMetArgProGluIleTTPiIleAlaGln 150  
 |||||  
 481 TGGAGATCTTACCAA.....GTGGCACA 506  
 |||||  
 151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyr 163  
 |||||  
 507 AATATCCAC.....GTGAT.....GCTGGTAC 530  
 |||||

seq\_name: gb\_est1:AI971169

seq\_documentation\_block:  
 LOCUS AI971169 492 bp mRNA linear EST 08-MAR-2000  
 DEFINITION wr24h12.x1 NCI\_CGAP\_Pr28 Homo sapiens cDNA clone IMAGE:2488679 3'  
 similar to TR:043522 043522 BML. [1] ;, mRNA sequence.

ACCESSION AI971169  
 VERSION AI971169  
 KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 492)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaaps-r@mail.nih.gov

Tissue procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 712 Std Error: 0.00

Seq primer: -400P from Gibco

High quality sequence stop: 450.

FEATURES

SOURCE

1..492

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2488679"

/clone\_lib="NCI\_CGAP\_Pr28"

/sex="male"

/dev\_stage="adult"

/lab\_host="DH10B"

/note="Organ: prostate; Vector: pTT30-Pac (Pharmacia)  
 with a modified polylinker; Plasmid DNA from the  
 normalized library NCI\_CGAP\_Pr22 was prepared, and ss  
 circles were made in vitro. Following HAP purification,  
 this DNA was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from a pool  
 of 5,000 clones made from the same library (cloneIDs  
 983608-986759, 1101192-1101959, and 1217928-1220615).  
 Subtraction by Bento Soares and M. Fatima Bonaudo."

BASE COUNT 119 a 107 c 130 g 134 t 2 others

ORIGIN

alignment\_scores:

Quality: 335.00 Length: 131

Ratio: 4.786 Caps: 1

Percent Similarity: 53.435 Percent Identity: 53.435

alignment\_block:

US-09-508-832-10 x AI971169/rev

Align seg 1/1 to reverse of: AI971169 from: 1 to: 492

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluClyAr 17  
 |||||  
 425 ATGGCAAGCAACCTTCTGATGTAAGTTCTTGAGTGCACCGAAGAGTAG 376  
 |||||

17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProt 34  
 |||||  
 375 ACAATTGCAGCTCGGAGAGGCTCCACAGCTCAGACCTGGGGCCCTA 326  
 |||||

34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50  
 |||||  
 325 CCTCCCTACACAGAGCCACAA..... 303  
 |||||

51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67  
 |||||  
 303 ..... 303

67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeta 84  
 |||||  
 303 ..... 303

84 rgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheasp 100  
 |||||  
 303 ..... 303

101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117  
 |||||  
 302 ...GACAGGAGCCAGCAGCCATGAGTTGTGCAAAATCAACAACNNC 256  
 |||||

117 oSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMet 131  
 |||||  
 255 AAGTCTCTTTCAGGAGCCCTCAACCACTATCTCAGTGCATG 213  
 |||||

seq\_name: gb\_est1:AW629314

seq\_documentation\_block:

LOCUS AW629314 664 bp mRNA linear EST 31-MAR-2000

DEFINITION h156602.x1 Soares\_NFL\_T\_GBC\_S1 Homo sapiens cDNA clone

IMAGE:2976314 3' similar to TR:043522 043522 BML. [1] ;, mRNA

sequence.

ACCESSION AW629314

VERSION AW629314.1 GI:7376104

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 664)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaaps-r@mail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -400P from Gibco

High quality sequence stop: 458.

FEATURES

source

1..664

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2976314"

/clone\_lib="Soares\_NFL\_T\_GBC\_S1"

/lab\_host="DH10B"

/note="Organ: pooled; Vector: pTT30-Pac (Pharmacia) with

a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;

Equal amounts of plasmid DNA from three normalized

libraries (fetal lung NBHL19W, testis NHT, and B-cell

NCI\_CGAP\_Gcbl) were mixed, and ss circles were made in

vitro. Following HAP purification, this DNA was used as

tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 176 a 131 c 148 g 208 t 1 others  
ORIGIN

## alignment\_scores:

Quality: 261.00 Length: 154  
Ratio: 3.575 Gaps: 3  
Percent Similarity: 47.403 Percent Identity: 39.610

## alignment\_block:

US-09-508-832-10 x AM629314/rev ..

Align seg 1/1 to reverse of: AM629314 from: 1 to: 664

```
3 LysGlnProSerAspValSerSerGluCysAspArgGluGlyArgGlnLe 19
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
664 AAGCCACCTTTTGATGAACCTTTGAGCTGCGCNGAGAGGTAGACAAAT 615

19 uGlnProAlaGluArgProGlnLeuArgProGlyAlaProThrSerL 36
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
614 GCAGGCTGCGAAGAGGCTTCCCGAGTTCAGACTTGGGGCCCCCTACCTCT 565

36 euGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGlyGly 52
:::|||||:|||||
564 ACAAGACAGACGCCACA..... 548

53 AspSerCysProHisGlySerProGlnGlyProLeuAlaProAlaSe 69
548 ..... 548

69 rProGlyProPheAlaThrArgSerProLeuPheIlePheMetArgArg 86
548 ..... 548

86 erSerLeuSerArgSerSerSerGlyTyPheSerPheAspThrAsp 102
||| |||
547 .....GAC 545

103 ArgSerProAlaProMetSerCysAspLysSerThrGlnThrProSerPr 119
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
544 AGGAGCCCGCCAGCCATGAGTTGTGCAAAATCAACACAAACCCCAAGTCC 495

119 oProCysGlnAlaPheAsnHisTyrLeuSerAlaMet..... 131
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
494 TCCTTGCCAGGCCCTCAACCACATCTCTAGTGTATGGATGACTCCGCTG 445

132 ..AlaSerMetArgGlnAlaGluProAlaAspMetArg...ProGluIle 146
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
444 GATCTCTCCTCAGAAATGCCCCCTCATAGGGAAGTTCAGTGGCCACTCAAG 395

147 TrpIleAlaGln 150
|||||:|||||:
394 TGGTTAGCAAAA 383
```